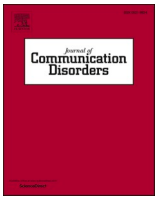
Journal of Communication Disorders 119 (2026) 106611

Contents lists available at ScienceDirect

Journal of Communication Disorders

journal homepage: www.elsevier.com/locate/jcomdis



The impact of exposure to additional languages and cognitive factors on narrative macrostructure in autistic and neurotypical children

Pauline Wolfer [a,*], Freideriki Tselekidou [b], Franziska Baumeister [a], Natalia Gagarina [b], Stephanie Durrleman [a]

[a] *Autism, Bilingualism, Cognitive and Communicative Development Research Group, Faculty of Science and Medicine, University of Fribourg, Switzerland*
[b] *Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Germany*

ARTICLE INFO

Keywords:
Narrative macrostructure
Autism spectrum disorder
Bilingualism
Theory of mind
Working memory
Metalinguistic awareness

ABSTRACT

This study adopts a twofold approach to examine how bilingual exposure and (socio-)cognitive factors influence narrative macrostructure in children. First, it investigated the impact of *Exposure to Additional Languages* (EtAL) on the narrative macrostructure of 90 autistic and 168 neurotypical children (ages 5–12), tested in their most proficient language, either English, French, (Swiss-) German, Italian or Spanish. Macrostructure was assessed through story structure, story complexity, and the production of Internal State Terms (IST), using a series of generalized mixed effects models controlling for age, sex, language skills, and non-verbal IQ. Among neurotypical children, greater EtAL predicted better story structure compared to peers with lower EtAL. No such effect was found in autistic children, suggesting that the potential benefits of EtAL may not extend equally across populations. Additionally, EtAL did not predict story complexity or IST-use in either group. Second, the study investigated the impact of Theory of Mind (ToM), working memory (WM), and metalinguistic awareness (MA) on macrostructure. While ToM and WM showed no significant effects, MA positively predicted story structure across groups. Findings indicate that dimensions of narrative macrostructure may benefit from exposure to multiple languages and MA, especially in neurotypical children. Results further support that, despite advice being provided to families of children with autism to abandon bilingualism, bilingual exposure is not detrimental to the communicative development of autistic children. Leveraging a large, neurodiverse sample and a continuous, hypothesis-driven measure of bilingualism, this study provides generalizable insights into how bilingualism and (socio-)cognitive factors influence narrative development across neurodiverse populations.

1. Introduction

Bilingualism, the use of and exposure to multiple languages regardless of proficiency (Grosjean, 1982), is known to influence various aspects of communication, language development, and cognition (Adesope et al., 2010; Bialystok, 2009; Bonuck et al., 2022). Research suggests that bilingual exposure influences linguistic outcomes, yet its specific effects on narrative macrostructure, the

* Corresponding author.
E-mail address: pauline.wolfer@unifr.ch (P. Wolfer).

https://doi.org/10.1016/j.jcomdis.2025.106611
Received 10 July 2025; Received in revised form 2 December 2025; Accepted 10 December 2025
Available online 11 December 2025
0021-9924/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

overall organization and structure of a story (Berman, 2009; Trabasso & Nickels, 1992), remains unclear. Exposure to additional languages (EtAL) provides access to diverse linguistic structures and storytelling conventions, potentially influencing how individuals structure and organize their narratives (i.e., narrative macrostructure) in the dominant language (Bruner, 1990; Minami, 2002). Different languages emphasize distinct ways of sequencing events, expressing causality, and expressing mental state (Bliss & McCabe, 2008; Tannen, 1980). As a result, children with greater EtAL may develop richer narratives than peers with lower EtAL, integrating elements from multiple languages. Yet, the impact of EtAL on macrostructure remains underexplored, especially in children with Autism Spectrum Disorder (ASD), who often face challenges in this realm (Baixauli et al., 2016).

Effective narrative production not only relies on rich language abilities, but also on higher-order (socio-)cognitive abilities, such as the ability to understand and attribute thoughts and emotions to others (Theory of Mind; ToM), to temporarily store and manipulate information (Working Memory; WM), and the ability to reflect on and manipulate language structures (Metalinguistic Awareness; MA). Autistic children, who often face challenges in these very domains, also experience difficulties with narrative coherence and organization. However, the relative contributions of ToM, WM, and MA to narrative macrostructure in both autistic and neurotypical (NT) children remain insufficiently understood.

The current study aimed to address these knowledge gaps by (1) examining whether EtAL impacted the narrative macrostructure of NT and autistic children, and (2) determining the respective contributions of ToM, EF and MA in narrative macrostructure. Specifically, we were interested in examining whether children with a greater exposure to languages other than their most proficient language (i.e., strongest language; S1) would present richer macrostructure skills. This continuous operationalization of bilingual exposure overcomes the limitations of monolingual *versus* bilingual group comparisons (De Houwer, 2023) and allows a more nuanced investigation of how varying levels of bilingual exposure impact narrative macrostructure in both autistic and NT children.

2. Narrative macrostructure

2.1. Narrative macrostructure: theoretical background

Oral narrative production, the ability to produce verbal "account[s] of events occurring over time" (Bruner, 1991, p. 6), is fundamental to human communication (Gagarina et al., 2016). Strongly linked to the development of literacy and associated with academic achievement (Bishop & Edmundson, 1987; Bliss et al., 1998), narratives develop from preschool throughout primary school (Berman, 2009). Narratives are typically analyzed at two distinct, yet complementary levels of competence: *microstructure*, the linguistic-related aspects of the discourse, such as morphosyntactic accuracy and lexical diversity, and (2) *macrostructure*, referring to the overarching organization, structure and coherence of the story, assumed to reflect "underlying mental narrative schema[s]" (Bohnacker, 2016, p. 21).

The richness of narrative productions can be analyzed through *story grammar models* (Stein & Glenn, 1979), a theoretical framework enabling researchers to examine macrostructural elements systematically across studies with various populations. According to this framework, stories contain both a *setting,* on the one hand, and an *episode* structure that relates the events, on the other (Stein & Glenn, 1979; Westby, 2005). While the setting specifies the context of the story (e.g., time and place, protagonists), the *episode* consists of a sequence of different constitutive components: (1) a Goal (G) pursued by the protagonist, (2) a protagonist's Attempt (A) to reach the goal, and (3) an Outcome (O) resulting from the attempt and pertaining to the goal (Westby, 2005).

This framework underpinned the *Multilingual Assessment Instrument for Narratives* (MAIN; Gagarina et al., 2012, 2019), a tool to assess narrative skills via telling, retelling and model story in various populations. MAIN comprises four parallel stories, each with a carefully designed six-picture sequence controlled for micro- and macrostructure, as well as linguistic, cognitive and cultural appropriateness. It additionally provides standardized scoring scheme to enable systematic evaluation. Because of this theoretically based design, standardized procedure of elicitation, scoring and analyses, MAIN has been adapted in over 100 languages and dialects and is widely used by researchers and clinicians in bilingual populations of various ages, cultural backgrounds and developmental profiles, such as children with Developmental Language Disorders (DLD), Down Syndrome or ASD (see Lindgren et al., 2023 for an overview; Mattiauda et al., 2024).

MAIN enables researchers to assess three macrostructural dimensions: *story structure, story complexity* and the *use of internal states terms. Story structure* indexes the quantity aspect, reflecting the number of episodic elements such as goals, attempts, outcomes stated by the speaker, while *story complexity* reflects the quality dimension, measuring the combination of these elements to deliver well-formed episodes (i.e., with the full Goal-Attempt-Outcome sequence showing higher complexity and more developed narrative construction; Gagarina et al., 2012). Finally, the use of *IST* reflects the speaker's "awareness of the goals and intentions of characters" (Altman et al., 2016, p. 168). *Internal state terms* refer to words describing thoughts, feelings, intentions and internal experiences of oneself and others, such as metacognitive verbs (e.g., to think, to want) or words expressing emotions (e.g., happy). On the one hand, IST provides an indication of a speaker's Theory of Mind (ToM), the ability to attribute mental states to others, and to predict their behaviors based on these beliefs (Baron-Cohen et al., 1985; Premack & Woodruff, 1978). On the other hand, they appear as relevant contributors to produce coherent and developed narratives, as a speaker must consider both the mental representation of the story's characters and the listener's state of knowledge to be rightfully informative (Altman et al., 2016; Westby, 2005). As such, narrative macrostructure involves different socio-cognitive factors, including ToM (see Section 4).

2.2. Narrative macrostructure in ASD

Children diagnosed with ASD experience persistent difficulties with social communication while presenting a great heterogeneity

of cognitive and linguistic profiles (American Psychiatric Association, 2022; Schaeffer et al., 2023; Silleresi, 2023). A meta-analysis conducted on studies investigating narrative skills in autistic children with an IQ in the normal range revealed major challenges in narrative macrostructure, as compared to neurotypical (NT) peers, with a large-effect size (Baixauli et al., 2016). Narrative productions of autistic children and adolescents tend to be less coherent, less well connected and structured within the *story grammar* framework than those of neurotypical peers (Baixauli et al., 2016; Diehl et al., 2006; Stirling et al., 2014). Children with ASD have been shown to produce narratives lacking key story grammar elements, and containing more irrelevant comments (Norbury et al., 2014). As effective storytelling also relies on socio-pragmatic skills, such as adapting to the context and the audience (Ketelaars et al., 2016; Norbury et al., 2014), difficulties in narrative macrostructure have been associated with their challenges with socio-pragmatic abilities (Beauchamp et al., 2023). Pertaining to IST-use, findings are rather mixed: While some studies find that autistic children produce fewer terms relative to mental and affective states than NT peers (Begeer et al., 2010; Diehl et al., 2006), other studies observe no differences in IST production between groups (Capps et al., 2000; Greco et al., 2023; Tager-Flusberg & Sullivan, 1995).

Narrative macrostructure thus varies across populations with different cognitive profiles such as autistic children, and is likewise influenced by diverse language experiences, including bilingualism.

3. Bilingualism effects

Bilingualism is inherently multidimensional, shaped by factors such as exposure, proficiency, and age of acquisition (De Houwer, 2023; Marian & Hayakawa, 2021; Surrain & Luk, 2019). Yet much of the literature has relied on a binary contrast between *"monolingual"* and *"bilingual,"* an oversimplification that may contribute to mixed findings.

3.1. Bilingualism effects on narrative macrostructure in NT children

Investigating the impact of bilingualism on narrative macrostructure provides valuable insights into how dual language exposure affects narrative skills. Studies on the relationship between bilingualism and narrative macrostructure in NT children have yielded controversial results. A handful of studies comparing mono- and bilingual children showed that bilinguals outperform monolinguals on narrative macrostructure (e.g., Chen & Yan, 2011; Tsimpli et al., 2016 on story structure and IST). Specifically, the study by Tsimpli et al. (2016) revealed that typically developing bilingual children performed better in story structure and produced significantly more ToM-related language, namely emotion and mental state terms, in their narratives than their monolingual peers. In contrast, the work by Lindgren (2018) demonstrated that Turkish-Swedish bilinguals exhibited notably lower performance in story structure compared to their Swedish monolingual peers. Furthermore, other studies on macrostructure production found comparable performance between monolinguals and bilinguals on story structure, IST use (Boerma et al., 2016; Haman et al., 2017; Rodina, 2017 for the second-language (L2) Norwegian of Russian-Norwegian bilinguals) and story complexity (Kunnari et al., 2016).

Apart from a monolingual *versus* bilingual categorization, some other studies have sought to highlight the effect of exposure on narrative macrostructure, generally focusing on bilingual children only. They opted for different measures, including the amount of exposure, the Age of Onset of bilingualism, or the Length of Exposure, and yielded mixed results (Lindgren et al., 2023). Particularly, in terms of amount of exposure, studies usually focused purely on whether the exposure to a particular language was crucial for narrating in that language (e.g., Lindgren, 2018). Evidence on whether exposure to additional languages impacts narrative production in the strongest language (S1) is limited. For instance, Haman et al (2017) investigated both heritage, first-acquired language (L1) and majority L2 language exposure on L1 narrative performance of Polish-English bilingual children. Results indicated that children's narrative skills in L1 were strongly associated with exposure both to the L1 and the L2. This evidence supports the idea that exposure to additional languages benefits the production of narrative macrostructure in bilinguals.

Differences in storytelling conventions across languages and cultures may also shape children's narrative production. Speakers may structure events and discourse elements in different ways depending on the languages they are exposed to (Bruner, 1990; Minami, 2002). Cultural communicative styles differ in the focus of the story (e.g., emphasis on *how* or *why* an event happened; Tannen, 1980), or on what is considered a 'good story' (chronological, goal-driven sequence versus overall thematic maintenance, indication of emotions; Bliss & McCabe, 2008; Minami, 2002; Tannen, 1980). Since narrative macrostructure is largely language-independent (Bohnacker et al., 2022; Lindgren et al., 2023), children may benefit from these different storytelling styles and integrate them in the narrative macrostructure of their S1, thereby resulting in richer narrative macrostructure. Although we do not directly assess cultural differences in this study, this perspective motivates examining exposure to additional languages as a potentially enriching factor.

3.2. Bilingualism effects on narrative macrostructure in autistic children

In research focusing on the autistic population, findings are scarcer, although more consistent. A handful of studies investigating bilingualism effects on narrative macrostructure reported null effects (Baldimtsi et al., 2016; Beauchamp et al., 2023; Hoang et al., 2018; Peristeri et al., 2020). Peristeri et al. (2020) examined narrative abilities in a sample of twenty bilingual and twenty monolingual autistic boys aged 7;5 - 12;0 matched on chronological age. While bilingual children outperformed their monolingual peers in story structure complexity – a composite measure including both the presence of story grammar elements (i.e., story structure) and the complexity of goal-directed episodes (i.e., story complexity), no group difference was observed in the production of IST. These findings closely replicated those of Baldimtsi et al. (2016), which used a similar narrative elicitation task and design in a smaller, independent sample of six bilingual and six monolingual autistic children of the same age range. Hoang et al. (2018) further confirmed the lack of

differences in the use of evaluative devices (including IST), in a study including five autistic bilinguals and five autistic monolinguals. In addition, groups did not differ in their measure of *story structure* either. Finally, Beauchamp et al. (2023) reported no difference in narrative story structure in 10-year-old autistic bilinguals, as compared to their aged-matched monolingual peers. In this study, relative exposure to the language of testing (in proportion) did not significantly correlate with narrative macrostructure in the autistic participants ($N = 134$; Beauchamp et al., 2023).

These studies offer insights into the narrative abilities of autistic children from various language backgrounds. They are valuable to the autistic community, where language-related challenges are often a primary concern for caregivers (Beauchamp & MacLeod, 2017; Drysdale et al., 2015). However, the limited sample sizes (i.e., six participants in Baldimtsi et al., (2016); five in Hoang et al., (2018); twenty in Peristeri et al., (2020) in each group) raises questions about the generalizability of the findings and the possibility to detect significant effects (Schäfer & Schwarz, 2019).

Both the heterogeneity observed in the NT population (Section 3.1), and the apparent absence of effect in the autistic population (Section 3.2), may reflect the oversimplified treatment of bilingualism as a binary construct (i.e., monolingual *versus* bilingual groups). Bilingualism is inherently multidimensional, with individuals varying notably in their language exposure, proficiency levels, age and mode of acquisition, language combinations (Marian & Hayakawa, 2021; Unsworth, 2016). Grouping participants based on an arbitrary bilingual criterion (e.g., a specific percentage of exposure to a second language) may obscure the nuanced individual differences in the bilingual experiences (De Houwer, 2023). Yet, measuring and accurately capturing bilingualism is challenging (Kaśćelan et al., 2021; Marian & Hayakawa, 2021; Surrain & Luk, 2019). Against this background, our study uses a continuous, hypothesis-driven measure of exposure to additional languages to test whether exposure predicts macrostructure in autism and to evaluate whether metalinguistic mechanisms (rather than exposure alone) better account for clinically meaningful aspects of narrative development (see also Marian & Hayakawa, 2021; Surrain & Luk, 2019; Hantman et al., 2023). Providing a transparent description of the sample and justifying the selection of bilingual variables based on research hypotheses may help disentangle the specific impact of targeted bilingual dimensions on the outcome (Byers-Heinlein et al., 2019), in both neurotypical (Surrain & Luk, 2019) and autistic individuals (Hantman et al., 2023). A hypothesis-driven approach to select the bilingual variables under study may also clarify which aspects of the bilingual experience are most relevant (Hantman et al., 2023; Surrain & Luk, 2019).

4. Cognitive factors related to narrative macrostructure

Delivering rich, coherent narratives arguably involves the interplay of several socio-cognitive factors besides mere linguistic skills. Specifically, it has been hypothesized that the lower performance in narrative macrostructure exhibited by autistic children could result from reduced abilities in skills involved in the narrative task, such as ToM and WM (Baixauli et al., 2016).

Theory of Mind plays an important role in effective storytelling by allowing individuals to convey story characters' mental states, thereby promoting detailed and cohesive narratives (Nicolopoulou & Richner, 2007; Pelletier & Astington, 2004; Trabasso & Nickels, 1992). Previous research reports positive associations between ToM abilities and narrative competence, in neurotypical children (Gamannossi & Pinto, 2014). Similarly, associations between ToM understanding and narrative competence are described for autistic individuals (Capps et al., 2000; Tager-Flusberg & Sullivan, 1995). However, inconsistent findings are reported in both populations (absence of correlation in NT children in Charman & Shmueli-Goetz, 1998; in ASD: Losh & Capps, 2003).

In addition, producing a coherent narrative requires higher-order skills such as WM. Working memory refers to the ability to temporarily hold information in mind and manipulate it in order to accomplish a task or purpose (Baddeley & Hitch, 1974). It is critical for reasoning, connecting thoughts and perceptions to create a meaningful knowledge, and it is a core aspect of executive functions, which are top-down cognitive processes implemented to achieve a goal (Diamond, 2013). Working memory is necessary for narrative construction, enabling the narrator to maintain a coherent overall structure by tracking the main events and characters, organizing them into a logical sequence, and ensuring consistency throughout the story (Veraksa et al., 2020; Whitely & Colozzo, 2013). Previous works showed positive associations between WM and narrative competence in NT children aged 5-6 (Veraksa et al., 2020; Whitely & Colozzo, 2013).

Finally, MA, the ability to reflect upon language as an object of thoughts (Tunmer & Herriman, 1984), also appears to be important for narrative macrostructure (Cortazzi & Jin, 2007; Kupersmitt et al., 2014; Russak & Zaretsky, 2021). For instance, MA may facilitate self-monitoring of speech, enabling speakers to detect and correct linguistic inaccuracies at the phonological, lexical or morpho-syntactic levels. This, in turn, supports narrative coherence and fluency, key components of macrostructure (Kormos, 2006; Levelt, 1989). It also helps refining lexical choices, to adjust and adapt to the audience and their feedback while telling the story (Cortazzi & Jin, 2007). Oral narratives elicited based on pictures require the understanding and synthesis of the events depicted into a coherent story so as to maintain the logical flow (Tunmer and Bowey, 1984). It requires understanding and inferring the temporal and causal relationships between the events for each protagonist, and organizing stories according to story schemata (Westby, 2005). Better knowledge of the story grammar elements that are present in any narratives, and awareness of the context of the story delivery, both lead to more coherent and well-formed narratives (Westby, 2005).

The assumption that autistic individuals would be impaired in narrative macrostructure because of primary difficulties in factors involved in the task is therefore plausible (Baixauli et al., 2016). Autistic children, indeed, present consistent impairments in both phonological and visual WM (Habib et al., 2019), and challenges with ToM are well established (e.g., Baron-Cohen et al., 1985; Mazza et al., 2017; Tager-Flusberg, 2007), compared to NT peers. In addition, autistic children have been reported to have difficulties with tasks engaging metalinguistic knowledge (Lewis et al., 2007), although recent work nuances this finding when language abilities are controlled for (Wolfer et al., 2024). While challenges with these skills may arguably have cascading effects on the narratives of autistic children, the empirical contribution of WM, ToM and MA on narrative macrostructure remains to be determined in children with

varying exposure to languages.

5. Research questions

The study has two main objectives. It first compared narrative macrostructure - measured as *story structure, story complexity*, and *IST-use* - in autistic and neurotypical children across several S1s: English, French, German, Swiss German, Italian, and Spanish. Cross-linguistic comparisons are justified given that story complexity has been shown to be relatively stable across languages and less dependent on language-specific features, while story structure, although more language-sensitive, can still be compared meaningfully when language background is considered (Gagarina et al., 2019; Bohnacker, 2016). In the present study, language was included as a covariate in the statistical models to account for both its impact and potential differences across S1s. Precisely, we investigated the impact of the exposure to additional languages over the past twelve months (EtAL) on narrative macrostructure, as well as the contribution of cognitive-related skills including WM, ToM, and MA. We specifically focused on current, relative exposure to additional languages over the past 12 months rather than categorizing bilinguals using a binary criterion (e.g., simultaneous vs sequential), so as to capture recent linguistic input and sensitivity to shifts in recent language environment. This time frame was chosen based on prior work showing that recent input is particularly relevant for language abilities in school-aged children (Beauchamp et al., 2023; Gonzalez-Barrero & Nadig, 2018; Unsworth, 2016). A 12-month window captures both stability and potential shifts in language environments (e.g., changes due to school, peers, or home use), and is commonly used in studies of bilingual language development. Given that language use and exposure fluctuate over time, measuring recent input provides valuable insight into how linguistic contact may influence narrative organization. Importantly, our research question concerns potential variation in the *expression* of macrostructural elements at school age rather than the *acquisition* of the macrostructural schema itself, which typically stabilises earlier in development (around ages 6-7; Berman & Slobin, 1994; Gagarina et al., 2019). Because languages differ in how macrostructural information is typically encoded, for instance in the explicitness of goals, causal links, or internal state terms (Bohnacker et al., 2022; Gagarina et al., 2019), children's current exposure determines the narrative models they are presently in contact with. Rather than focusing on cumulative exposure, which reflects long-term language history and may overestimate early input that no longer corresponds to children's present linguistic environments (e.g., following changes in schooling, peers, or family language practices), we focus on current exposure because it provides a more accurate representation of the models shaping children's narrative expression at the time of testing. Nonetheless, for completeness, cumulative exposure was also examined in SM7, and results are available for interested readers.

This study aimed to determine: *(1) How exposure to additional languages in the past twelve months (EtAL) affects the narrative macrostructure, measured as story structure, story complexity and IST, in the S1 of autistic and NT children?*

We predicted that, in both groups, increased EtAL would enhance macrostructure, measured as *story structure, story complexity and IST*. Prior research has shown that exposure to multiple languages can influence how individuals structure and organize narratives in their dominant language, by providing access to diverse linguistic structures, discourse styles, and storytelling conventions (Bruner, 1990; Minami, 2002). Therefore, greater EtAL was expected to enhance children's ability to organize, elaborate, and enrich their storytelling by providing linguistic and cultural richness and exposing them to multiple narrative styles and communicative conventions. This, in turn, may foster the ability to structure stories more effectively, incorporate nuanced details such as including internal state terms, resulting in richer narratives. Additionally, the exposure and managing of multiple linguistic systems may enhance children's understanding of story grammar elements, and the ability to monitor and sequence information to produce clear narratives. Alternatively, in line with prior research reporting null effects of bilingualism in the autistic population on macrostructure (e.g., Beauchamp et al., 2023; Hoang et al., 2018), the effect of EtAL may be absent in autistic children.

Since narrative macrostructure has been hypothesized to rely on ToM, WM and MA, we furthermore explored whether these factors would moderate the potential relationship between EtAL and narrative macrostructure in autistic and neurotypical children:

The second research question aimed to elucidate: *(2) How do ToM, WM and MA impact narrative macrostructure measured as story structure, story complexity and IST in autistic and NT children presenting different EtAL?*

We predicted that cognitive-related skills (i.e., ToM, WM, and MA) would positively impact narrative macrostructure in both neurotypical and autistic children. Specifically, for *story structure*, we predicted that greater ToM abilities, WM and MA would lead to higher production of story grammar elements. For *story complexity*, we predicted that ToM, WM and MA would significantly contribute to the production of more complex story episodes, since this dimension, and especially the production of its inferred components (e.g., Goal), is related to higher-order cognitive skills. It requires the ability to understand and express the protagonists' emotional state, but also to hold in memory and properly combine the story components (i.e., Goal, Attempt, Outcome) to achieve higher complexity of the story. For *IST*, we predicted that children with greater ToM would produce more IST, particularly terms related to ToM (i.e., ToM-related IST). We thus expected autistic children to produce fewer ToM-related IST than their neurotypical peers (Begeer et al., 2010; Diehl et al., 2006).

6. Methods

6.1. Participants

A total of $N = 258$ children (39% girls, 61% boys) between 5;0 and 11;11 years old participated in this study. The group of autistic participants included $N = 90$ children (106 $\pm$ 22.7 months; 5;1 – 11;11 years old), formally diagnosed with ASD following the criteria of the DSM-5 (American Psychiatric Association, 2022), prior to the study by a certified clinician. The group of neurotypical children

included $N = 168$ children (99.7 ± 23.3 months, 5;0 – 11;11) with no reported suspicion or diagnosis of ASD, nor any other developmental disorder. Caregivers completed the *Social Communication Questionnaire* (*SCQ*; Rutter et al., 2003), and NT participants had scores below 15, confirming the absence of significant social communication difficulties associated with ASD (see Supplementary Material SM1 for the SCQ score distribution across groups). Participants were recruited via flyers, websites, contact with clinicians, schools and autism centers. Children were tested in Europe (Switzerland, France, Germany, Spain, UK) and North America (USA, Canada).

NT children were significantly younger than autistic peers ($W = 8686$, $p = .049$), had stronger language skills in their most proficient language S1 ($W = 3847$, $p < .001$) and came from families with slightly higher parental education levels ($W = 4793$, $p < .001$), a reliable indicator of socioeconomic status (Hauser & Warren, 1997). Yet, parental education was high in both groups, with most caregivers holding post-secondary degrees in both groups (caregivers achieved either post-graduate education or hold a university degree; see Table 1). The autistic group included more boys (82%) than the NT group (49%), reflecting the well-documented sex ratio imbalance in autism(American Psychiatric Association, 2022; Zeidan et al., 2022). Following prior studies on narrative macrostructure in ASD (Hoang et al., 2018; Peristeri et al., 2020), only children with non-verbal IQ ≥ 80 were included. Autistic participants had significantly lower non-verbal IQ scores than NT peers ($W = 5503$, $p < .001$). Autistic participants had on average lower IQ than NT children ($W = 5503$, $p < .001$). Given group differences in variables known to influence narrative macrostructure - *age* (Bohnacker, 2016; Gagarina et al., 2016), non-verbal IQ (Botting, 2005), *S1 language skills* (Bohnacker, 2016; Fiani et al., 2022), and *sex* (Kauschke et al., 2016) - these were statistically controlled in statistical analyses. See Table 1 for detailed sample characteristics.

All children were tested in their most proficient language (i.e., strongest language, S1), defined via parental ratings from the Quantifying Bilingual EXperience questionnaire (*Q-BEx*; De Cat et al., 2022), either in English, French, German, Swiss German, Italian,

Table 1
Sample characteristics.

	Autistic children N = 90	**NT children N = 168**	**Test and *p*-value**
Age in months	106 (22.7)	99.7 (24.2)	$W = 8686$, $p = .048$
Median [Range]	111 [71 - 143]	97.5 [60 - 143]	
Sex assigned at birth			
Female	16 (17.8%)	85 (50.6%)	$\chi^2(1) = 25.16$, $p < .001$
Male	74 (82.2%)	83 (49.4%)	
Parental educational level[a]	4.01 (1.09)	4.70 (0.67)	$W = 4793$, $p < .001$
Median [Range]	4 [1 - 5]	5 [2 - 5]	
General Language skills[b]	-0.68 (1.63)	0.39 (0.95)	$W = 3847$, $p < .001$
Median [Range]	-0.75 [-4.0 - 1.70]	0.50 [-2.40 - 2.40]	
Non-verbal IQ	97.7 (12.6)	103 (11.0)	$W = 5503$, $p < .001$
Median [Range]	95.5 [80 - 133]	103 [82 - 133]	
SCQ score	20.1 (7.40)	4.20 (3.03)	$W = 5345$, $p < .001$
Median [Range]	20 [3 - 34]	4 [0 - 14]	
WM span	3.40 (2.23)	4.22 (1.75)	$W = 5560$, $p = .009$
Median [Range]	4 [0 - 6]	5 [0 - 6]	
ToM	12.7 (3.88)	17.8 (2.09)	$W = 1689.5$, $p < .001$
Median [Range]	12.7 [1.73 - 19.9]	18.6 [10.9 - 20.0]	
MA	0.77 (0.26)	0.84 (0.24)	$W = 5629$, $p = .005$
Median [Range]	0.86 [0 - 1]	0.93 [0 - 1]	
Testing language			
N English	11 (12.2%)	33 (19.6%)	
Age (months), Mean (SD)	112 (26.2)	98.4 (26.6)	
Median [Range]	[70 - 140]	[61 - 143]	
N French	34 (37.8%)	38 (22.6%)	
Age (months), Mean (SD)	102 (24.1)	94.9 (21.7)	
Median [Range]	[61 - 141]	[61 - 138]	
N German	16 (17.8%)	56 (33.3%)	
Age (months), Mean (SD)	107 (20.4)	103 (21.5)	
Median [Range]	[74 - 143]	[60 - 139]	
N Swiss German	8 (8.9%)	16 (9.5%)	
Age (months), Mean (SD)	107 (20.4)	112 (22.9)	
Median [Range]	[64 - 128]	[77 - 138]	
N Italian	2 (2.2%)	14 (8.3%)	
Age (months), Mean (SD)	88.0 (32.5)	83.6 (27.1)	
Median [Range]	[65 - 111]	[63 - 136]	
N Spanish	19 (21.1%)	11 (6.5%)	
Age (months), Mean (SD)	111 (20.6)	107 (25.4)	
Median [Range]	[80 - 143]	[82 - 141]	

Note. Unless specified, Mean (SD) are presented. SCQ : Social Communication Questionnaire (range 0 – 40). Missing data: SCQ (ASD = 6, NT = 2); WM (ASD = 5, NT = 5); ToM (ASD = 3, NT = 2); MA (ASD = 4, NT = 1).

[a] Parental education was determined by the highest reported value between the two caregivers, corresponding to the highest level of schooling achieved, specifically (1) primary school, (2) middle school, (3) high school, (4) post-graduate education, and (5) university.

[b] Composite z-score for receptive vocabulary and morphosyntax.

or Spanish. This questionnaire was used to retrieve information of the participants' experience on up to three languages, such as their relative exposure and estimated proficiency in each language. For each language, parents rated the child's ability in *speaking, understanding, reading,* and *writing* using a 4-point Likert scale (0 = hardly at all, 1 = not very well, 2 = pretty well, 3 = very well). An additional option, not applicable (NA), was available when a skill was not yet relevant (e.g., writing in pre-literate children). For each language, the valid subdomain scores were summed and averaged, excluding NA responses, and converted into a percentage of the maximum possible score. This approach generated an individual proficiency score for each language, capturing a comprehensive and developmentally appropriate overview of the child's skills without penalizing for age-typical gaps (e.g., not yet reading). The S1 corresponded to the language in which the participant demonstrated the highest proficiency. In cases of equal proficiency across languages, the testing language was selected using a stepwise procedure: first, protocol availability; then, child preference; and if no preference was expressed, the societal language used to communicate with the experimenter. EtAL was calculated as the percentage of time children were exposed to any language other than their S1 over the past 12 months. For example, a monolingual child had an EtAL of 0; a child exposed equally to S1 and L2 had an EtAL of 50.

In terms of broader bilingual experience, the sample included both monolingual and multilingual children. In the autistic group, 23% were monolingual, 51% bilingual, and 26% trilingual; in the neurotypical group, 29% were monolingual, 47% bilingual, and 24% trilingual. Most children were exposed to their most proficient language S1 from birth (66% of autistic children, 71% of neurotypical children), with the majority of the remainder acquiring it before age three. Only a small minority (6% autistic, 4% neurotypical) acquired their S1 after this age. For the second-best language L2, the median age of first exposure was 0 months in both groups, reflecting that many participants were simultaneous bilinguals, while others were first exposed later through school or community contexts. L3 exposure, when present, typically began later, specifically for autistic children (median: 36 months in autistic children; 1 month in neurotypical children). These figures highlight both the diversity of language backgrounds represented and the predominance of early exposure within the sample. Further details about the linguistic experience of the participants, such as age of first exposure, place of exposure, current and cumulative exposure, richness of input, rated proficiency and diversity of languages, are provided in Supplementary Material SM1.

6.2. Procedure

Participants were tested individually in a quiet room in their most proficient language. They were accompanied by a trained experimenter and could take as many pauses as they wanted. Testing took place in several sessions and included the tasks presented in the *Materials* section below. Full details of the tasks procedure, including examples and instructions, is presented on the OSF replication package, under https://osf.io/6xmj2/?view_only=cedcacf1d09c4906be2908d641efaa09.

6.3. Materials

6.3.1. Narrative assessment

Administration. The *Multilingual Assessment Instrument for Narratives* (*MAIN*; Gagarina et al., 2012, 2019) was used to elicit oral narratives in children's most proficient language. It was selected for its suitability with multilingual populations and potential for systematic comparisons across studies. The standardized task administration procedure and scoring scheme in "telling" mode were followed, and children were allowed to choose which story they wanted to tell, either *Baby Birds* or *Baby Goats*. These two parallel stories are considered comparable in terms of story length, story grammar, plotline and number of protagonists. Each story consisted of six pictures, depicting a series of events constituting three episodes (Goal – Attempt - Outcome). Participants were invited to look at the pictures as long as they wished. They were then asked to tell the story to a stuffed toy, a naive listener blind to the story who could not see the pictures. Pictures were unfolded two by two as the participant told the story. Forty-seven percent of the sample selected the *Baby Birds* story (respectively 51% of the autistic group, 43% of the NT group), 54% told the *Baby Goats* story (49% of the autistic group, 57% in the NT group).

Coding and Scoring. Productions were transcribed verbatim by native speakers of each language, namely the first author and trained research assistants blinded to the groups and study hypotheses. The MAIN scoring grid (Gagarina et al., 2019) was then followed to compute macrostructure scores: First, to assess *story structure (*i.e., the quantity measure), each stated element of the story grammar was awarded 1 point (i.e., story setting specification in time or place; IST statement to initiate an event or describe the final situation of an episode; (G) Goal, (A) Attempt or (O) Outcome mention). The maximum score possible was 17. Second, *story complexity* (i.e., the quality measure) was evaluated by determining the level of complexity in each of the three episodes. A score of 0 was given if no G, A or O was expressed in an episode. A score of 1 would be given if either A, O or AO was produced, a score of 2 would be given if either G, GA or GO was expressed. Finally, a score of 3 was given if a full GAO episode was verbalized. The maximum score for story complexity was 9 (max = 3 for each episode). Third, the production of *IST* was identified. Following Peristeri et al. (2020), IST were further classified into two separate categories: terms associated with ToM (i.e, called "ToM-related ISTs"), such as terms conveying emotions, intentions, desires (e.g., "sad", "think", "want"), and terms considered unrelated to ToM, such as perceptual, physiological or linguistic terms (i.e., "ToM-unrelated ISTs", like "see", "smell", "say"). Each occurrence, even repetitions, was awarded 1 point.

Inter-rater reliability. 26.79% of the transcriptions and coding were independently transcribed and coded a second time by native speakers of the respective languages. Coders were trained research assistants, blind to the study hypotheses and groups. Any discrepancy between the transcriptions and scores were discussed and resolved by mutual agreement to reach 100% similarity (inter-rater reliability before agreement: 83.10%).

6.3.2. Non-verbal IQ

The short, online version of the *Raven's Progressive Matrices* (Raven's-2; Raven et al., 2018) was used to assess non-verbal reasoning. Participants chose one option among five that best matched the geometric pattern containing a missing piece displayed on the screen. The resulting standardized non-verbal IQ score (100 ± 15) was computed.

6.3.3. Language skills

Participants completed a digital version of the *Test for Reception of Grammar* (TROG; Bishop, 2003) to assess their receptive morphosyntactic abilities in the testing language, and the digital version of the *Peabody Picture Vocabulary Test* (PPVT-4; Dunn & Dunn, 2007) evaluated their receptive vocabulary breadth (see Supplementary Material SM2 for the references of the different language versions). Participants listened to pre-recorded words (for the PPVT) and sentences with increasing morphosyntactic complexity (for the TROG) and selected the corresponding image from four options on screen. Z-scores, compared to the manuals' reference norms of each language, were calculated separately for the TROG and the PPVT. The average of the two z-scores was used in the analyses to create a composite score reflecting both lexical and morphosyntactic proficiency, following the approach of Engberg-Pedersen (2018) and Yau et al. (2016). Because no Swiss German version of the PPVT and TROG exists, the 16 participants (14 NT children, 2 autistic children) who told the story in Swiss German completed the German version of these tasks, as the two languages are closely related (Scherrer, 2012).

6.3.4. Theory of mind

Caregivers completed the *Theory of Mind Inventory 2* (ToMI-2; Hutchins et al., 2012) rating 60 statements on a 0–20 scale to assess children's everyday ToM abilities. The ToMI-2 has been validated in both typically developing and clinical populations, including autistic children, and demonstrates strong psychometric properties, with high internal consistency (Cronbach's $\alpha > .90$) and good construct validity (Hutchins et al., 2012; Hutchins & Prelock, 2018). It is widely used in research as an ecologically valid alternative to direct experimental tasks, as it captures a broad range of ToM-related behaviors observable in daily life. For this study, we focused on 16 items relevant to narrative production: the six assessing mental state understanding (e.g., *"My child understands the word 'want'."*) and the ten items targeting emotion recognition ("*My child recognizes when others are sad."*). The average of these items (range: 0–20) was used as the ToM score, reflecting socio-cognitive skills likely involved in the MAIN storytelling task. Language-appropriate adaptations of the ToMI-2 were used for all testing languages, ensuring comparable caregiver reports across the multilingual sample. All items were completed for participants included in the final analyses.

6.3.5. Working memory

To assess WM, the *Frogs Matrices Task* (FMT) adapted by Morales et al. (2013) was administered. Children had to recall the reverse order of positions of 2 to 6 frogs presented sequentially. The memory span was determined by the longest sequence of frogs that the participant could accurately recall in at least one of the two trials (i.e., ranging between 0 and 6; All participants included in the analyses completed the full task without missing trials).

6.3.6. Metalinguistic awareness

Metalinguistic awareness was assessed using a non-verbal, digital Grammatical Judgment Task (GJT) minimizing cognitive and linguistic demands (Wolfer et al., 2024). Participants judged the grammaticality of 16 semantically appropriate sentences (8 grammatically correct, 8 incorrect) by selecting either a candy or a sock to feed a cartoon monster. To account for response bias (e.g., tendency to judge all sentences as correct), an MA-score was calculated using the following formula from Linebarger et al. (1983): MA-score = 0.5 + (y – x) (1 + y – x) / 4y (1 x), where x = false alarm rate, and y = hit rate. Scores range from 0 (poor discrimination) to 1 (perfect), with 0.5 indicating chance performance. No trials were missing for participants included in the analyses. Full task procedures and materials are available in Wolfer et al. (2024) and in the OSF replication package https://osf.io/6xmj2/.

6.4. Statistical plan

To investigate the impact of *EtAL* on narrative macrostructure (RQ1), and the influence of WM, ToM, and MA on performance (RQ2), generalized linear models were fitted separately for each macrostructure dimension (i.e., *story structure, story complexity, IST-use*). Both RQ1 and RQ2 were examined with the same model for each dimension. All models included *age, group*, and *EtAL* (and their interactions), along with covariates (*language skills, non-verbal IQ,* and *sex*). Socio-cognitive variables were added stepwise, and model selection was based on AIC. Full information is presented in Supplementary Material SM3.

7. Results

Story Structure. On a possible maximum score of 17, the group of autistic children obtained a mean score of 5.42 ± 2.48 SD, ranging from 0 to 12 (median value: 5). The group of NT children obtained a mean score of 6.72 ± 2.16 SD, ranging from 1 to 13 (median value: 6.5).

The best model for story structure was the model with the maximal structure (Supplementary Material SM3, Table SM4A): it revealed a significant interaction between *group* and *EtAL* ($\beta = -0.20$, SE = 0.10, $z = -1.96$, $p = .049$). Simple effects analyses were further conducted to locate the source of the interaction. In NT children, EtAL positively impacted performance ($\beta = 0.11$, SE = 0.06, $z = 1.97$, $p = .048$; Table SM4B): NT children with greater EtAL obtained significantly better scores on story structure, than children with

lower EtAL (Fig. 1). This effect, however, was not detected in autistic children (β = -0.06, SE = 0.09, z = -0.71, p = .47; Table SM4C), suggesting a different impact of EtAL in the two groups.

Performance improved with age (β = 0.19, SE = 0.06, z = 2.85, p = .004; Table SM4A), and girls outperformed boys (β = 0.34, SE = 0.10, z = 3.47, p < .001). Metalinguistic awareness also significantly predicted story structure (β = 0.11, SE = 0.06, z = 1.98, p = .047), while ToM and WM did not yield significant effects (all ps > .05).

Story complexity. Autistic children obtained a mean score of 3.11 ± 1.28, ranging from 0 to 7 (median value: 3). Neurotypical children exhibited a mean score of 3.68 ± 1.22 SD, ranging from 1 to 7 (median value: 4).

The best model for story complexity was, again, the model with maximal structure (Supplementary Material SM3, Table SM5). It showed no significant main effects of *EtAL* (p > .05). Furthermore, the three-way interaction between *EtAL, age*, and *group* was also non-significant (β = 0.197, SE = 0.189, t = 1.043, p = .298), indicating no differential impact of EtAL on story complexity across groups or age. Similarly, the other cognitive and linguistic predictors did not reach significance (all ps > .05). Fig. 2 illustrates performance on story complexity as a function of EtAL in both groups.

IST-use. The group of autistic children produced an average of 2.34 ± 2.02 IST (1.16 ± 1.28 ToM-unrelated IST; 1.19 ± 1.48 ToM-related IST). Most of them produced 2 words; scores ranged from 0 to 7 IST produced. The group of NT children produced an average of 3.0 ± 2.13 IST (1.70 ± 1.53 ToM-unrelated IST; 1.30 ± 1.30 ToM-related IST). Scored ranged from 0 to 11, and most of the neurotypical children produced 3 IST.

The best model for IST-use was also the model with maximal structure (Supplementary Material SM3, Table SM6A). It revealed a significant three-way interaction between *EtAL, age* and *group* (β = 0.38, SE = 0.19, z = 1.98, p = .048). Post-hoc investigations revealed that *age* and *EtAL* did significantly impact the production of IST in the autistic group (β = -0.31, SE = 0.11, z = -2.76, p = .006, Table SM6B), but not in the neurotypical group (β = 0.08, SE = 0.05, z = 1.47, p = .14, Table SM6C). As shown in Fig. 3, autistic children displayed a differential impact of EtAL depending on age: At younger ages, those with lower EtAL tended to produce fewer ISTs than peers with greater EtAL. However, this pattern reversed at older ages, where children with lower EtAL showed a steeper increase in IST production. Nevertheless, the high variability in predicted performance suggests caution in interpreting this trend, particularly for older children, where confidence intervals widen considerably. Moreover, overall IST production remained relatively low across participants, suggesting that, despite the observed interaction, the practical impact of EtAL on IST production in autistic children may be limited. In contrast, in NT children (Fig. 3, right panel), the relationship between age and IST production remained stable across different levels of EtAL, indicating no meaningful interaction between these factors in this group.

EtAL significantly interacted with condition (β = 0.23, SE = 0.10, z = 2.33, p = .02): For both groups, ToM-related IST production decreased with greater EtAL (β = -0.20, SE = 0.08, z = - 2.57, p = .01; Table SM6D), while ToM-unrelated ISTs showed a non-significant increase (β = 0.02, SE = 0.06, p = .71; Table SM6E).

The interaction between *age* and *condition* was significant (β = 0.36, SE = 0.10, z = 3.72, p < .001; Table SM6A). These variables further interacted with *condition* (β = 0.38, SE = 0.22, z = 0.88, p = .0048). Specifically, among autistic children, older age was associated with fewer ToM-related ISTs (Table SM5F) and more ToM-unrelated ISTs, though these trends did not reach significance. No

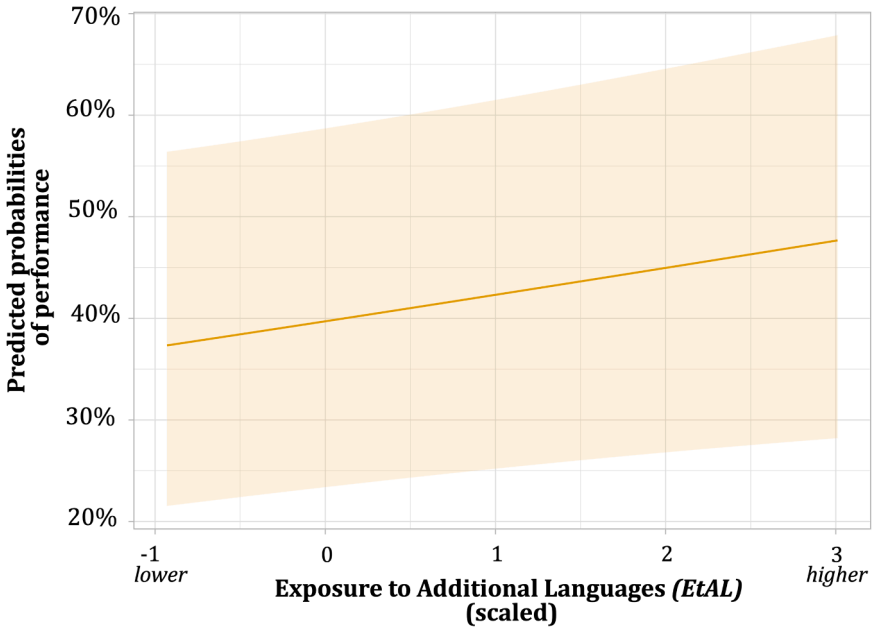


Fig. 1. Predicted story structure performance as a function of EtAL (scaled) in the group of NT children. Higher EtAL values indicate greater relative exposure to languages other than the child's L1 in the past 12 months. The solid line represents model predictions, and the shaded area shows 95% confidence intervals.

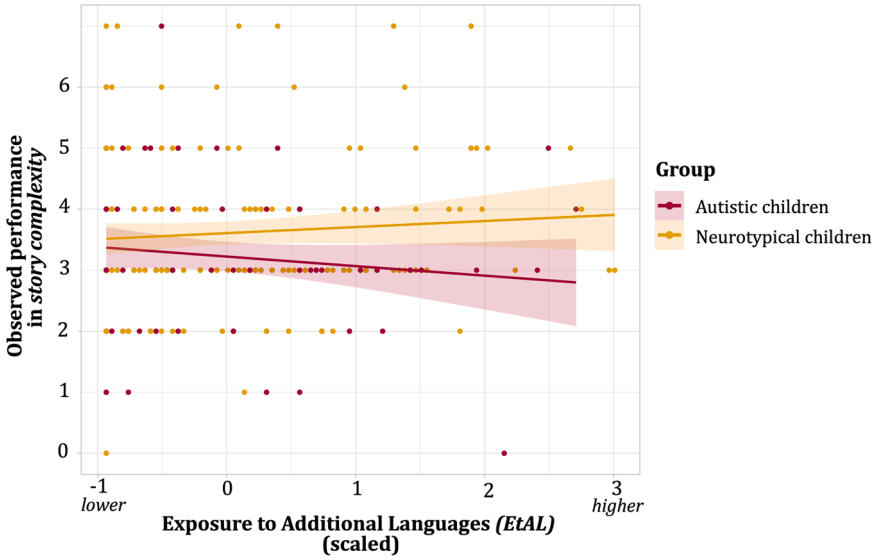


Fig. 2. Observed performance in *story complexity* as a function of EtAL (scaled), in autistic (red) and NT (yellow) children. Dots represent individual participants, solid lines show model predictions, and shaded areas show 95% confidence intervals.

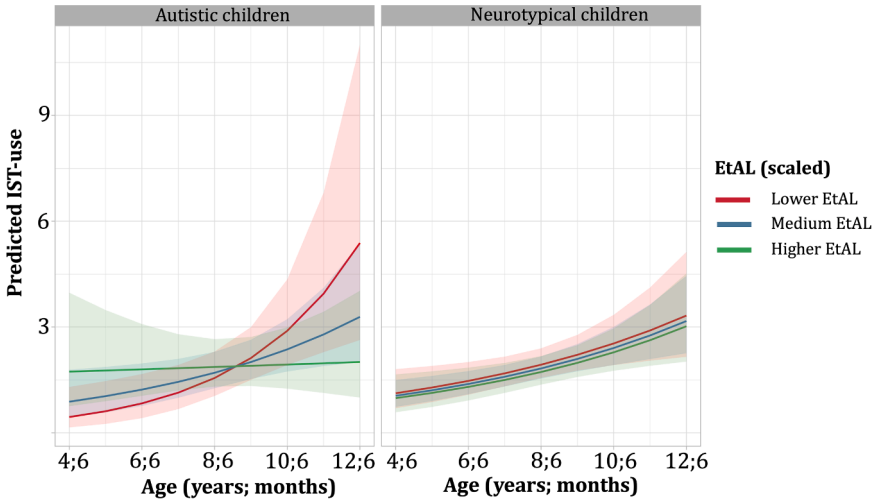


Fig. 3. Predicted IST-use as a function of age and EtAL (scaled), presented separately for autistic children (left panel) and NT children (right panel). Lines represent model-predicted values for lower (red), medium (blue), and higher (green) EtAL levels, with shaded areas showing 95% confidence intervals. Age is displayed in years; months to ease interpretation.

age-related differences were observed in NT children (Table SM6C).

Girls produced significantly more ISTs than boys ($\beta = 0.45$, SE = 0.09, z = 4.93, $p < .001$). Although ToM, WM, and MA improved model fit, none significantly predicted IST use (all *p*s $< .5$).

8. Discussion

This study examined the impact of the *Exposure to additional languages* (*EtAL)* on the narrative macrostructure in S1 of autistic and age-matched NT children between age 5 and 12. Accounting for age, sex, language skills, and non-verbal IQ, NT children benefitted from greater EtAL on their story structure. Specifically, neurotypical children with greater EtAL produced more elaborate narratives than their peers with lower EtAL, as reflected by the inclusion of more story structure elements (e.g., mention of the setting, a goal, attempt or outcome in an episode). This finding is consistent with the prediction that increased EtAL would support narrative macrostructure by fostering the inclusion of more relevant *story grammar* elements in children's S1 narratives. For instance, multilingual exposure may provide access to more diverse storytelling conventions and communicative practices, promoting a richer

understanding of how narratives are organized and presented across contexts (Bruner, 1990; Minami, 2002). This aligns with previous research reporting higher scores on story structure in the first-acquired language L1 for children with greater exposure to L1 and L2, than children with lower exposure to these languages (Haman et al., 2017). It also suggests that neurotypical children may take advantage of their exposure to additional languages to enrich their storytelling skills in their S1.

In contrast, we found no evidence that EtAL influenced story structure in autistic children, suggesting a different impact of EtAL between autistic and NT children. The absence of detectable EtAL effects in autistic children with varying EtAL aligns with previous research comparing monolingual and bilingual groups (Beauchamp et al., 2023; Hoang et al., 2018). One possible explanation for this difference is that autistic children process and generalize linguistic information differently than their NT peers: Difficulties with pragmatic inferencing and extracting implicit meaning from linguistic input (Bodner et al., 2015; Loukusa et al., 2007) could limit their ability to identify and apply relevant storytelling elements across languages. Additionally, challenges in generalizing acquired knowledge may further constrain their ability to transfer storytelling skills from their additional languages to S1, preventing the macrostructural advantages observed in neurotypical children (Carruthers et al., 2020; de Marchena et al., 2015). Importantly, the finding of no detrimental effect of EtAL on the storytelling abilities of autistic children supports accumulating evidence that autistic children with dual-language exposure may achieve similar language and communicative outcomes than their monolingually-raised peers (Beauchamp et al., 2020; Hambly & Fombonne, 2012; Peristeri et al., 2020).

As for *story complexity,* no effect of EtAL was found in any of the groups. While this finding may appear at odds with studies reporting a bilingual advantage in *story (structure) complexity* (e.g., Baldimtsi et al., 2016; Peristeri et al., 2020), several methodological differences are important to consider. These studies used a composite measure termed *story structure complexity*, which combined both the presence of story structure elements and episodic elaboration. In contrast, the present study assessed these components separately. Moreover, previous studies relied on categorical group comparisons (bilingual vs. monolingual), which may obscure individual variability in language experience (De Houwer, 2023). By treating exposure to additional languages as a continuous variable, the present approach offers a nuanced perspective on how a specific aspect of the bilingual experience may influence linguistic outcomes.

The absence of impact of EtAL on *story complexity* aligns with previous findings suggesting the relative language-independence of the skill to combine elements in narratives. This higher-level of narrative organization is cognitively more demanding than story structure. It is possible that the cognitive and linguistic benefits of additional language exposure take longer to manifest in complex storytelling abilities. Since story complexity involves global coherence, integration of multiple elements, and hierarchical structuring, it may require sustained and long-term bilingual experience to observe developmental differences. In contrast, story structure may be more immediately influenced by language exposure, particularly through frequent engagement with diverse linguistic input. In summary, if a storyteller is able to understand the content of picture stimuli and convey the core meaning of a narrative episode, this ability reflects cognitive maturation rather than language proficiency *per se* (Gagarina et al., 2016). The present findings contribute additional evidence in support of a multidimensional model of narrative organization, highlighting the relative independence of this skill from EtAL.

EtAL impacted the production of IST in different ways. First, in autistic children, but not in neurotypical children, EtAL interacted with age in predicting IST: At a younger age, autistic children with greater EtAL produced more ISTs than those with lower EtAL. However, this trend reversed at older ages, where children with lower EtAL showed greater IST-use. Yet, due to the high variability and the continuous nature of the predictors, further examination of these tendencies was not possible. This finding suggests that the role of EtAL in IST-use (1) differs between autistic and neurotypical children, and (2) may change over time in autism. Second, findings revealed that as EtAL increased, the production of ToM-related ISTs decreased. This was unexpected, as bilingualism has been associated with enhanced ToM abilities (e.g., Adesope et al., 2010; Kovács, 2009). One possible explanation is that the use of ISTs, particularly those related to ToM, relies heavily on lexical knowledge, situating them more within the domain of microstructure than macrostructure. Therefore, greater exposure to additional languages may not facilitate IST use unless it is accompanied by sufficient vocabulary development in the language of testing. Alternatively, it is possible that a longer duration of multilingual exposure, rather than current EtAL is necessary for effects on ToM-related ISTs to emerge.

At the same time, the interaction between EtAL and age in autistic children, and the additional interaction between *age, group*, and *condition* (ToM-related, ToM-unrelated ISTs) suggests that these patterns may also reflect underlying socio-cognitive and communicative difficulties associated with autism, rather than a direct effect of bilingual exposure. While NT children showed a general increase in both ToM-related and ToM-unrelated ISTs with age, autistic children tended to produce fewer ToM-related ISTs but more ToM-unrelated ISTs as they grew older. This pattern may reflect difficulties in integrating mental-state reasoning into narratives, aligning with broader challenges in social cognition and communication in autism (American Psychiatric Association, 2022; Boucher, 2012).

Furthermore, there was a significant positive effect of age on *story structure*, aligning with previous studies (Bohnacker, 2016; Bohnacker et al., 2022; Gagarina et al., 2015). Girls significantly exhibited better story structure and produced more ISTs than boys, consistent with the literature (Kauschke et al., 2016).

The second objective was to investigate the impact of cognitive narrative-related factors (ToM, WM and MA) on macrostructure. ToM is relevant because successful storytelling requires the narrator to represent and convey characters' mental states, intentions, and emotions. Narratives that include such perspective-taking are typically more cohesive, as events are linked not only through time but also through characters' goals and beliefs. Children with more advanced ToM therefore tend to produce more integrated stories (Fernández, 2013). WM also plays a central role, since narrators must hold multiple elements of a story in mind while planning and articulating them. Effective storytelling depends on keeping track of characters, settings, and events while updating the sequence as the plot unfolds. Children with stronger WM capacities are better able to manage this cognitive load, resulting in more coherent and structurally complex narratives (Khan, 2013). MA further supports narrative production by enabling children to reflect on language

itself and apply story grammar deliberately, which helps to organize episodes into a coherent whole. Taken together, these socio-cognitive skills provide complementary support for the development of narrative macrostructure.

While these factors significantly improved model fit for all three macrostructure dimensions evaluated (*story structure, story complexity, IST-use*), only MA emerged as a significant independent positive predictor of story structure. This suggests that greater ability to reflect on and organize linguistic content at a higher level (Bialystok, 2001; Karmiloff-Smith, 1986) enhances macrostructure performance, particularly in terms of story structure. This is because the skill to produce macrostructure requires not only recognition of the sequences of events, but also encoding and manipulation of the linguistic forms to make a coherent whole. MA equips children with this skill to consciously attend to and manipulate linguistic forms and this helps to organize story structures into this coherent whole.

Interestingly, while ToM and WM did not independently predict any of the three macrostructure outcomes, their inclusion in the model nonetheless reduced unexplained variance, improving overall model precision and suggesting potential indirect contributions to narrative production.

Moreover, this ability to reflect on and manipulate language itself may directly support the organization of coherent stories, whereas ToM and WM may play more indirect or limited roles.

One possible explanation for these null effects lies in the specificity of the measures used. For ToM, we relied on a ToMI-2 subscore reflecting children's understanding of mental state vocabulary (e.g., "know," "want"), which was chosen for its relevance to narrative production. However, narrative performance in the MAIN requires not only understanding mental state vocabulary but also the ability to observe, interpret, and integrate characters' emotions and intentions in context, particularly for earning points on story structure and IST-use. These situational inferencing demands may tap somewhat different aspects of ToM than those measured here. Similarly, our WM measure focused on non-verbal span. While non-verbal WM has been linked to narrative ability (Veraksa et al., 2020; Whitely & Colozzo, 2013), prior studies indicate that the association with narrative skills is typically stronger for verbal WM than for visual WM (Veraksa et al., 2020). Thus, the absence of significant effects may reflect the specificity of the constructs measured, rather than the absence of a relationship.

A further explanation may relate to developmental factors. Macrostructural narrative skills typically reach a developmental plateau around age 6–7 (Berman & Slobin, 1994; Gagarina et al., 2019), and many children in our sample may have already achieved this level of narrative organization. Moreover, both WM and ToM show the most rapid development in early childhood (up to around age 6–7), with more gradual refinement continuing through middle childhood (Gathercole et al., 2004; Wellman, 2014). Previous research has shown that WM contributes to narrative abilities in younger children, but its role may diminish as narrative skills become more established (Friend & Bates, 2014; Scionti et al., 2023; Veraksa et al., 2020). It is therefore plausible that in this age range, limited developmental variability in these domains reduced their predictive power. The effect of EtAL persisted even after accounting for ToM and WM, suggesting additional mechanisms at play. Future research is warranted to confirm these possibilities.

Taken together with prior studies showing null or inconsistent bilingual effects in autism (Baldimtsi et al., 2016; Beauchamp et al., 2023; Hoang et al., 2018; Peristeri et al., 2020), our results suggest that while EtAL enriches story structure in NT children, in autism exposure alone is unlikely to yield this effect. Clinical progress may require explicit supports, such as training in mental-state and causal language, and elicitation beyond visual retells to open-ended and autobiographical forms (Botting et al., 2010; Hilvert et al., 2016; King & Palikara, 2018; Marini et al., 2020; Masoumi et al., 2024; Siller et al., 2014; Tam et al., 2022; Westby, 2022). More broadly, the findings highlight metalinguistic awareness as a promising target for intervention across groups. Strengthening children's ability to reflect on and apply story grammar, for example through explicit instruction, or cross-linguistic comparisons, may enhance story structure in both autistic and neurotypical children.

8.1. Limitations

Due to the nature of the MAIN task, only children with verbal abilities and sufficient attentional cognitive resources could participate in this study. As a result, autistic children with intellectual and/or language impairment were not represented, reflecting a common methodological selection bias (Russell et al., 2019). Additionally, this study focused on the impact of bilingual exposure on macrostructure but did not examine other dimensions of bilingual experience (e.g., age of acquisition of the language, richness of the input received in a language, linguistic distance between language pairs, etc.), despite their potential influence on narrative macrostructure (Lindgren et al., 2023).

Moreover, while our models included random intercepts for participants (and items where applicable), we did not include testing language or site as random effects. This decision was guided by the standardized nature of the MAIN protocol, designed for multilingual research, and by the fact that all testing languages were Western European and broadly comparable in narrative macrostructure. Because of this, our findings should be interpreted within a Western European linguistic and cultural context. Nonetheless, subgroup sizes did not permit direct modeling of language effects, and potential variation across languages or cultural storytelling conventions cannot be fully excluded, even though individual proficiency in the testing language was statistically controlled. Future research with larger samples per language and from more diverse language families will be needed to address this more directly.

Additionally, while we noted that storytelling conventions can vary across languages and cultures (Bruner, 1990; Minami, 2002; Bliss & McCabe, 2008), we did not measure such variation directly. Although the MAIN protocol minimizes language-specific influences and ensures comparability across testing languages, future research could investigate more explicitly how exposure to diverse cultural storytelling traditions may contribute to children's narrative development. Future work should elicit narratives in both heritage/home and school languages, model dominance, balance, and context of use, as well as culturally adapt coding (Manolitsi & Botting, 2011; Peristeri et al., 2017). Recruiting across socioeconomic strata and storytelling traditions, documenting home literacy

practices, and sampling diverse language families, with discourse-relational analyses of causal/intentional links, will strengthen ecological validity. Broader elicitation beyond retells to include open-ended and autobiographical narratives, plus outcomes such as teacher/peer ratings and repair frequency, ideally with longitudinal follow-up, would capture communicative impact more directly (King & Palikara, 2018; Wantzen et al., 2021; Westby, 2022). Integrating mechanism-sensitive measures, such as intention-focused ToM, nonverbal temporal sequencing, and planning/monitoring, could clarify when EtAL and metalinguistic awareness translate into macrostructural and communicative gains (Carlsson et al., 2020; Lee et al., 2018; Marini et al., 2020; Siller et al., 2014).

Finally, in the present study, we adopted a categorical distinction between autistic and neurotypical children, based on prior diagnosis, while we used the SCQ measure only to confirm the group assignment. Although this approach aligns with our research questions and much of the existing literature on narrative macrostructure, future research could consider adopting a continuous approach to autistic traits (e.g., Kašćelan et al., 2019) to provide additional insight into how variability in symptom severity impacts narrative abilities.

8.2. Conclusion

This study investigated how EtAL in the past 12 months impacted the narrative macrostructure of school-aged autistic and NT children. Participants told a story elicited from pictures in their most proficient language. Findings revealed that NT children with greater EtAL presented richer narratives in terms of *story structure* than peers with lower EtAL. This suggests that children with exposure to multiple languages may draw from this exposure to enrich their storytelling skills. This effect, however, was not present in autistic children, where EtAL did not affect macrostructure. Consequently, while children with ASD may not benefit from EtAL to the same extent as their NT peers, bilingual exposure was not found to be detrimental to their narrative macrostructure abilities. This finding is particularly relevant from a societal perspective, as families of autistic children are often advised to restrict their child to a single language due to concerns that bilingualism may hinder language development (Howard et al., 2021; Yu, 2013). Our results contribute to a growing body of evidence demonstrating that autistic children, like their neurotypical peers, can develop successfully in multilingual environments (Drysdale et al., 2015; Gilhuber et al., 2023).

Data availability statement

The data that support the findings of this study, and code to reproduce the analyses are openly available in OSF at https://osf.io/6xmj2/.

Funding details

This study is supported by the Swiss National Science Foundation (SNSF) awarded to Stephanie Durrleman (grant PR00P1_193104/1).

Compliance with ethical standards and informed consent

This study was approved by the Swiss Association of Research Ethics Committees *Swissethics* (Project ID-2022-00878), the Institutional Review Board of the University of Connecticut (US), the Research Ethics Board Office at McGill University in Montreal (Canada), the Psychology Research Ethics Committee of the University of Edinburgh (UK), the Institutional Review Board of Emerson College Boston (US), and the Research Ethics Committee of the Unio Catalana Hospitals Foundation (Spain). All parents provided informed written consent for their child's participation prior to their inclusion in the study. All work has been carried out in accordance with The World Medical Association's Declaration of Helsinki for experiments involving humans.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT to assist with revision and proofreading. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Pauline Wolfer: Writing – original draft, Visualization, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Freideriki Tselekidou:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Franziska Baumeister:** Writing – review & editing, Supervision, Resources, Methodology, Investigation, Conceptualization. **Natalia Gagarina:** Writing – review & editing, Conceptualization. **Stephanie Durrleman:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors report there are no competing interests to declare.

Acknowledgments

We are very grateful to all the families for their participation in the study. Thanks to all experimenters for their investment in data collection and to the students and interns for the transcriptions and scoring of the narratives in the different languages. Many thanks to Yulia Zuban for her careful reading of the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jcomdis.2025.106611.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207–245. https://doi.org/10.3102/0034654310368803

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

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders (DSM-5-TR)* (5th Edition). text rev.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation, 8* pp. 47–89). Elsevier. https://doi.org/10.1016/S0079-7421(08)60452-1

Baixauli, I., Colomer, C., Roselló, 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, & D. Waughtal (Eds.), *Proceedings of the 40th annual Boston university conference on language development* (pp. 18–31). Cascadilla Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition, 21*(1), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

Beauchamp, M. L. H., & MacLeod, A. A. N. (2017). Bilingualism in children with autism spectrum disorder: Making evidence based recommendations. *Canadian Psychology, 58*(3), 250–262. https://doi.org/10.1037/cap0000122

Beauchamp, M. L. H., Rezzonico, S., Bennett, T., Duku, E., Georgiades, S., Kerns, C., Mirenda, P., Richard, A., Smith, I. M., Szatmari, P., Vaillancourt, T., Waddell, C., Zaidman-Zait, A., Zwaigenbaum, L., & Elsabbagh, M. (2023). The influence of bilingual language exposure on the narrative, social and pragmatic abilities of school-aged children on the autism spectrum. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-022-05678-w

Beauchamp, M. L. H., Rezzonico, S., & MacLeod, A. A. N. (2020). Bilingualism in school-aged children with ASD: A pilot study. *Journal of Autism and Developmental Disorders, 50*(12), 4433–4448. https://doi.org/10.1007/s10803-020-04501-8

Begeer, S., Malle, B. F., Nieuwland, M. S., & Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: Comparing individuals with high-functioning autism and typically developing controls. *European Journal of Developmental Psychology, 7*(1), 104–122. https://doi.org/10.1080/17405620903024263

Berman, R. A. (2009). *Language development in narrative contexts*. The Cambridge Handbook of Child Language.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study, 748*. Lawrence Erlbaum Associates, Inc. xiv.

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics, 21*, 169–181.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3–11. https://doi.org/10.1017/S1366728908003477

Bishop, D. V. (2003). *Test for reception of grammar, version 2. London: Pearson assessment*. Addison-Wesley Professional.

Bishop, D. V., & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech & Hearing Disorders, 52*(2), 156–173. https://doi.org/10.1044/jshd.5202.156

Bliss, L. S., & McCabe, A. (2008). Personal narratives: Cultural differences and clinical implications. *Topics in Language Disorders, 28*(2), 162–177. https://doi.org/10.1097/01.TLD.0000318936.31677.2d

Bliss, L. S., Mccabe, A., & Miranda, A. E. (1998). Narrative assessment profile: Discourse analysis for school-age children. *Journal of Communication Disorders, 31*(4), 347–363. https://doi.org/10.1016/s0021-9924(98)00009-4

Bodner, K. E., Engelhardt, C. R., Minshew, N. J., & Williams, D. L. (2015). Making inferences: Comprehension of physical causality, intentionality, and emotions in discourse by high-functioning older children, adolescents, and adults with autism. *Journal of Autism and Developmental Disorders, 45*(9), 2721–2733. https://doi.org/10.1007/s10803-015-2436-3

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/10.1111/1460-6984.12234

Bohnacker, U. (2016). Tell me a story in English or Swedish: Narrative production and comprehension in bilingual preschoolers and first graders. *Applied Psycholinguistics, 37*(1), 19–48. https://doi.org/10.1017/S0142716415000405

Bohnacker, U., Lindgren, J., & Öztekin, B. (2022). Storytelling in bilingual Turkish-Swedish children: Effects of language, age and exposure on narrative macrostructure. *Linguistic Approaches to Bilingualism, 12*(4), 413–445. https://doi.org/10.1075/lab.20057.boh

Bonuck, K., Shafer, V., Battino, R., Valicenti-McDermott, R. M., Sussman, E. S., & McGrath, K. (2022). Language disorders research on bilingualism, school-age, and related difficulties: A scoping review of descriptive studies. *Academic Pediatrics, 22*(4), 518–525. https://doi.org/10.1016/j.acap.2021.12.002

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry, 46*(3), 317–326. https://doi.org/10.1111/j.1469-7610.2004.00355.x

Botting, N., Riches, N., Gaynor, M., & Morgan, G. (2010). Gesture production and comprehension in children with specific language impairment. *British Journal of Developmental Psychology, 28*(1), 51–69. https://doi.org/10.1348/026151009×482642

Boucher, J. (2012). Research review: Structural language in autistic spectrum disorder – characteristics and causes. *Journal of Child Psychology and Psychiatry, 53*(3), 219–233. https://doi.org/10.1111/j.1469-7610.2011.02508.x

Bruner, J. (1990). *Acts of meaning, 179*. Harvard University Press. xvii.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry, 18*(1), 1–21.

Byers-Heinlein, K., Esposito, A. G., Winsler, A., Marian, V., Castro, D. C., & Luk, G. (2019). The case for measuring and reporting bilingualism in developmental research. *Collabra: Psychology, 5*(1), 37. https://doi.org/10.1525/collabra.233

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. https://doi.org/10.1023/a:1005126916531

Carlsson, E., Åsberg Johnels, J., Gillberg, C., & Miniscalco, C. (2020). Narrative skills in primary school children with autism in relation to language and nonverbal temporal sequencing. *Journal of Psycholinguistic Research, 49*(3), 475–489. https://doi.org/10.1007/s10936-020-09703-w

Carruthers, S., Pickles, A., Slonims, V., Howlin, P., & Charman, T. (2020). Beyond intervention into daily life: A systematic review of generalisation following social communication interventions for young children with autism. *Autism Research, 13*(4), 506–522. https://doi.org/10.1002/aur.2264

Charman, T., & Shmueli-Goetz, Y. (1998). The relationship between theory of mind, language and narrative discourse: An experimental study. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 17*(2), 245–271.

Chen, L., & Yan, R. (2011). Development and use of English evaluative expressions in narratives of Chinese–English bilinguals. *Bilingualism: Language and Cognition, 14*(4), 570–578.

Cortazzi, M., & Jin, L. (2007). Narrative learning, EAL and metacognitive development. *Early Child Development and Care, 177*(6-7), 645–660. https://doi.org/10.1080/03004430701379074

De Cat, C., Kašćelan, D., Prevost, P., Serratrice, L., Tuller, L., & Unsworth, S. (2022). *Quantifying Bilingual EXperience (Q-BEx): Questionnaire manual and documentation*. https://osf.io/v7ec8/.

De Houwer, A. (2023). The danger of bilingual–monolingual comparisons in applied psycholinguistic research. *Applied Psycholinguistics, 44*(3), 343–357. https://doi.org/10.1017/S014271642200042X

de Marchena, A. B., Eigsti, I.-M., & Yerys, B. E. (2015). Brief report: Generalization weaknesses in verbally fluent children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(10), 3370–3376. https://doi.org/10.1007/s10803-015-2478-6

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diehl, J. J., Bennetto, L., & Young, E. C. (2006). Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *Journal of Abnormal Child Psychology, 34*(1), 83–98. https://doi.org/10.1007/s10802-005-9003-x

Drysdale, H., van der Meer, L., & Kagohara, D. (2015). Children with autism spectrum disorder from bilingual families: A systematic review. *Review Journal of Autism and Developmental Disorders, 2*(1), 26–38. https://doi.org/10.1007/s40489-014-0032-7

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.

Engberg-Pedersen, E. (2018). Inference from academic texts in children with autism. *Pragmatics & Cognition, 25*(2), 363–383. https://doi.org/10.1075/pc.18011.eng

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

Fiani, R., Henry, G., & Prévost, P. (2022). Macrostructure in narratives produced by Lebanese arabic-French bilingual children: Developmental trends and links with language dominance, exposure to narratives and lexical skills. *Linguistic Approaches to Bilingualism, 12*(4), 446–478. https://doi.org/10.1075/lab.20035.fia

Friend, M., & Bates, R. P. (2014). The union of narrative and executive function: Different but complementary. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00469

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2012). MAIN: Multilingual assessment instrument for narratives. *ZAS Papers in Linguistics, 56*. https://doi.org/10.21248/zaspil.56.2019.414, 155–155.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2015). 9. Assessment of narrative abilities in bilingual children. In S. Armon-Lotem, J. de Jong, & & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 243–276). Multilingual Matters. https://www.degruyterbrill.com/document/doi/10.21832/9781783093137-011/html?srsltid=AfmBOoq0H-RJDI-LQDD_F_u0cIPEgiYutwQhc1xdaELjmCvIK6vtZln2.

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

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

Gamannossi, B. A., & Pinto, G. (2014). Theory of mind and language of mind in narratives: Developmental trends from kindergarten to primary school. *First Language, 34*(3), 262–272. https://doi.org/10.1177/0142723714535875

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gilhuber, C. S., Raulston, T. J., & Galley, K. (2023). Language and communication skills in multilingual children on the autism spectrum: A systematic review. *Autism: The International Journal of Research and Practice, 27*(6), 1516–1531. https://doi.org/10.1177/13623613221147780

Gonzalez-Barrero, A. M., & Nadig, A. (2018). Bilingual children with autism spectrum disorders: The impact of amount of language exposure on vocabulary and morphological skills at school age. *Autism Research, 11*(12), 1667–1678. https://doi.org/10.1002/aur.2023

Greco, G., Choi, B., Michel, K., & Faja, S. (2023). Here's the story: Narrative ability and executive function in autism spectrum disorder. *Research in Autism Spectrum Disorders, 101*, Article 102092. https://doi.org/10.1016/j.rasd.2022.102092

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Harvard University Press.

Habib, A., Harris, L., Pollick, F., & Melville, C. (2019). A meta-analysis of working memory in individuals with autism spectrum disorders. *PLoS ONE, 14*(4), Article e0216198. https://doi.org/10.1371/journal.pone.0216198

Haman, E., Wodniecka, Z., Marecka, M., Szewczyk, J., Białecka-Pikul, M., Otwinowska, A., Mieszkowska, K., Łuniewska, M., Kołak, J., Miękisz, A., Kacprzak, A., Banasik, N., & Foryś-Nogala, M. (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? *Evidence from Polish-English Migrants to the United Kingdom. Frontiers in Psychology, 8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01444.

Hambly, C., & Fombonne, E. (2012). The impact of bilingual environments on language development in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(7), 1342–1352. https://doi.org/10.1007/s10803-011-1365-z

Hantman, R. M., Choi, B., Hartwick, K., Nadler, Z., & Luk, G. (2023). A systematic review of bilingual experiences, labels, and descriptions in autism spectrum disorder research. *Frontiers in Psychology, 14*, Article 1095164.

Hauser, R. M., & Warren, J. R. (1997). Socioeconomic indexes for occupations: A review, update, and critique. *Sociological Methodology, 27*(1), 177–298. https://doi.org/10.1111/1467-9531.271028

Hilvert, E., Davidson, D., & Gámez, P. B. (2016). Examination of script and non-script based narrative retellings in children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 29-30*, 79–92. https://doi.org/10.1016/j.rasd.2016.06.002

Hoang, H., Gonzalez-Barrero, A. M., & Nadig, A. (2018). Narrative skills of bilingual children with autism spectrum disorder. *Discours. Revue de Linguistique, Psycholinguistique et Informatique. A Journal of Linguistics, Psycholinguistics and Computational Linguistics, 23*, 23. https://doi.org/10.4000/discours.9856. Article.

Howard, K., Gibson, J., & Katsos, N. (2021). Parental perceptions and decisions regarding maintaining bilingualism in autism. *Journal of Autism and Developmental Disorders, 51*(1), 179–192. https://doi.org/10.1007/s10803-020-04528-x

Hutchins, T., & Prelock, P. (2018). Using Story-Based interventions to improve episodic memory in autism spectrum disorder. *Seminars in Speech and Language, 39*(02), 125–143. https://doi.org/10.1055/s-0038-1628365

Hutchins, T. L., Prelock, P. A., & Bonazinga, L. (2012). Psychometric evaluation of the Theory of Mind Inventory (ToMI): A study of typically developing children and children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 42*(3), 327–341. https://doi.org/10.1007/s10803-011-1244-7

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition, 23*(2), 95–147. https://doi.org/10.1016/0010-0277(86)90040-5

Kašćelan, D., Katsos, N., & Gibson, J. L. (2019). Relations between bilingualism and autistic-like traits in a general population sample of primary school children. *Journal of Autism and Developmental Disorders, 49*(6), 2509–2523. https://doi.org/10.1007/s10803-019-03994-2

Kašćelan, D., Prévost, P., Serratrice, L., Tuller, L., Unsworth, S., & De Cat, C. (2021). A review of questionnaires quantifying bilingual experience in children: Do they document the same constructs? *Bilingualism: Language and Cognition*, 1–13. https://doi.org/10.1017/S1366728921000390

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*(3), 840–852. https://doi.org/10.1007/s10803-015-2620-5

Ketelaars, M. P., Jansonius, K., Cuperus, J., & Verhoeven, L. (2016). Narrative competence in children with pragmatic language impairment: A longitudinal study. *International Journal of Language & Communication Disorders, 51*(2), 162–173. https://doi.org/10.1111/1460-6984.12195

Khan, K. S. (2013). *The Relation Between Cognitive Skills and Language Skills in Typically-developing 3½-to-6 Year Olds* [Ph.D.]. https://www.proquest.com/docview/1774432360/abstract/610E0F7D3E634515PQ/1.

King, D., & Palikara, O. (2018). Assessing language skills in adolescents with autism spectrum disorder. *Child Language Teaching and Therapy, 34*(2), 101–113. https://doi.org/10.1177/0265659018780968

Kormos, J. (2006). *Speech production and second language acquisition* (p. 221). Lawrence Erlbaum Associates Publishers. xxvii.

Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science, 12*(1), 48–54. https://doi.org/10.1111/j.1467-7687.2008.00742.x

Kunnari, S., Välimaa, T., & Laukkanen-Nevala, P. (2016). Macrostructure in the narratives of monolingual Finnish and bilingual Finnish–Swedish children. *Applied Psycholinguistics, 37*(1), 123–144. https://doi.org/10.1017/S0142716415000442

Kupersmitt, J., Yifat, R., & Blum-Kulka, S. (2014). The development of coherence and cohesion in monolingual and sequential bilingual children's narratives: Same or different? *Narrative Inquiry, 24*(1), 40–76. https://doi.org/10.1075/ni.24.1.03kup

Lee, M., Martin, G. E., Hogan, A., Hano, D., Gordon, P. C., & Losh, M. (2018). What's the story? A computational analysis of narrative competence in autism. *Autism: The International Journal of Research and Practice, 22*(3), 335–344. https://doi.org/10.1177/1362361316677957

Levelt, W. J. M. (1989). *Speaking: From intention to articulation, 566*. The MIT Press. xiv.

Lewis, F. M., Murdoch, B. E., & Woodyatt, G. C. (2007). Communicative competence and metalinguistic ability: Performance by children and adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*(8), 1525–1538.

Lindgren, J. (2018). *Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish Children aged 4–6 [Phd]*. Acta Universitatis Upsaliensis.

Lindgren, J., Tselekidou, F., & Gagarina, N. (2023). Acquisition of narrative macrostructure: A comprehensive overview of results from the multilingual assessment instrument for narratives. *ZAS Papers in Linguistics, 65*, 111–132. https://doi.org/10.21248/zaspil.65.2023.623

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. https://doi.org/10.1023/A:1024446215446

Loukusa, S., Leinonen, E., Kuusikko, S., Jussila, K., Mattila, M.-L., Ryder, N., Ebeling, H., & Moilanen, I. (2007). Use of context in pragmatic language comprehension by children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 37*(6), 1049–1059. https://doi.org/10.1007/s10803-006-0247-2

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. https://doi.org/10.1177/0265659010369991

Marian, V., & Hayakawa, S. (2021). Measuring bilingualism: The quest for a "bilingualism quotient". *Applied Psycholinguistics, 42*(Suppl 2), 527–548. https://doi.org/10.1017/s0142716420000533

Marini, A., Ozbič, M., Magni, R., & Valeri, G. (2020). Toward a definition of the linguistic profile of children with autism spectrum disorder. *Frontiers in Psychology, 11*, 808. https://doi.org/10.3389/fpsyg.2020.00808

Masoumi, E., Soleymani, Z., Modarresi, Y., Maroufizadeh, S., & Mohseni, F. (2024). The effect of narrative intervention on spoken narrative and reading comprehension abilities in Farsi-speaking children with autism spectrum disorders: A randomized clinical trial. *Medical Journal of the Islamic Republic of Iran, 38*, 118. https://doi.org/10.47176/mjiri.38.118

Mattiauda, E., Özsoy, O., Gagarina, N., & Perovic, A. (2024). Inferential abilities in down syndrome: Examining verbal and nonverbal contributors to narrative comprehension in adolescents and adults. In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *46. Proceedings of the annual meeting of the cognitive science society* (pp. 3804–3811). https://escholarship.org/uc/item/6fr5k0c2.

Mazza, M., Mariano, M., Peretti, S., Masedu, F., Pino, M. C., & Valenti, M. (2017). The role of theory of mind on social information processing in children with autism spectrum disorders: A mediation analysis. *Journal of Autism and Developmental Disorders, 47*(5), 1369–1379. https://doi.org/10.1007/s10803-017-3069-5

Minami, M. (2002). *Culture-specific language styles: The development of oral narrative and literacy*. Multilingual Matters.

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology, 114*(2), 187–202. https://doi.org/10.1016/j.jecp.2012.09.002

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

Norbury, C. F., 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

Pelletier, J., & Astington, J. W. (2004). Action, consciousness and theory of mind: Children's ability to coordinate story characters' actions and thoughts. *Early Education and Development, 15*(1), 5–22. https://doi.org/10.1207/s15566935eed1501_1

Peristeri, E., Andreou, M., & Tsimpli, I. M. (2017). Syntactic and story structure complexity in the narratives of high- and low-language ability children with autism spectrum disorder. *Frontiers in Psychology, 8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02027.

Peristeri, E., Baldimtsi, E., Andreou, 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*. https://doi.org/10.1016/j.jcomdis.2020.105999

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526. https://doi.org/10.1017/S0140525×00076512

Raven, J. C., Rust, J., Chan, F., & Zhou, X. (2018). *Raven's progressive matrices 2, Clinical Edition* (Pearson).

Rodina, Y. (2017). Narrative abilities of preschool bilingual Norwegian-Russian children. *International Journal of Bilingualism, 21*(5), 617–635. https://doi.org/10.1177/1367006916643528

Russak, S., & Zaretsky, E. (2021). Cognitive and linguistic skills associated with cross-linguistic transfer in the production of oral narratives in English as a foreign language by Arabic- and Hebrew-speaking children: Finding common denominators. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.664152

Russell, G., Mandy, W., Elliott, D., White, R., Pittwood, T., & Ford, T. (2019). Selection bias on intellectual ability in autism research: A cross-sectional review and meta-analysis. *Molecular Autism, 10*, 9. https://doi.org/10.1186/s13229-019-0260-x

Rutter, M., Bailey, A., Berument, S. K., & Lord, C. (2003). *Social communication questionnaire: Manual*. Western Psychological Services.

Schaeffer, J., Abd El-Raziq, M., Castroviejo, E., Durrleman, S., Ferré, S., Grama, I., Hendriks, P., Kissine, M., Manenti, M., Marinis, T., Meir, N., Novogrodsky, R., Perovic, A., Panzeri, F., Silleresi, S., Sukenik, N., Vicente, A., Zebib, R., Prévost, P., & Tuller, L. (2023). Language in autism: Domains, profiles and co-occurring conditions. *Journal of Neural Transmission*. https://doi.org/10.1007/s00702-023-02592-y

Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology, 10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00813.

Scherrer, Y. (2012). *Generating swiss German sentences from standard German: A multi-dialectal approach* [Université de Genève]. https://archive-ouverte.unige.ch/unige:26361.

Scionti, N., Zampini, L., & Marzocchi, G. M. (2023). The relationship between narrative skills and executive functions across childhood: A systematic review and meta-analysis. *Children, 10*(8), 8. https://doi.org/10.3390/children10081391. Article.

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

Silleresi, S. (2023). *Developmental profiles in autism spectrum disorder*. John Benjamins Publishing Company.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children: A test of a schema. In R. O. Freedle (Ed.), *New Directions in Discourse Processing* (pp. 53–120). Ablex Publishing Company.

Stirling, L., Douglas, S., Leekam, S., & Carey, L. (2014). Chapter 8. The use of narrative in studying communication in Autism Spectrum Disorders: A review of methodologies and findings. In J. Arciuli, & J. Brock (Eds.), *Communication in Autism* (pp. 171–216). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.11.09sti.

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition, 22*(2), 401–415. https://doi.org/10.1017/S1366728917000682

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science, 16*(6), 311–315. https://doi.org/10.1111/j.1467-8721.2007.00527.x

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*(3), 241–256. https://doi.org/10.1017/S0142716400007281

Tam, J. Y. C., Barrett, E. A., & Ho, A. Y. H. (2022). Narrative intervention in school-aged children with autism spectrum disorder: A systematic review. *Speech, Language and Hearing, 25*(4), 463–480. https://doi.org/10.1080/2050571X.2021.1985893

Tannen, D. (1980). A comparative analysis of oral narrative strategies: Athenian Greek and American English. In W. L. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 51–87). Ablex Publishing Company.

Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes, 15*(3), 249–275. https://doi.org/10.1080/01638539209544812

Tsimpli, I. M., Peristeri, E., & Andreou, M. (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguistics, 37*(1), 195–216. https://doi.org/10.1017/S0142716415000478

Tunmer, W. E., & Bowey, J. A. (1984). Metalinguistic awareness and reading acquisition. In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic awareness in children: Theory, research, and implications* (pp. 144–168). Springer. https://doi.org/10.1007/978-3-642-69113-3_9.

Tunmer, W. E., & Herriman, M. L. (1984). The development of metalinguistic awareness: A conceptual overview. In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic awareness in children: Theory, research, and implications* (pp. 12–35). Springer. https://doi.org/10.1007/978-3-642-69113-3_2.

Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. *6. Quantity and quality of language input in bilingual language development* (pp. 103–122). De Gruyter Mouton. https://doi.org/10.1515/9783110341249-008

Veraksa, A., Bukhalenkova, D., Kartushina, N., & Oshchepkova, E. (2020). The relationship between executive functions and language production in 5–6-year-old children: Insights from working memory and storytelling. *Behavioral Sciences, 10*(2), 2. https://doi.org/10.3390/bs10020052. Article.

Wantzen, P., Boursette, A., Zante, E., Mioche, J., Eustache, F., Guénolé, F., Baleyte, J.-M., & Guillery-Girard, B. (2021). Autobiographical memory and social identity in autism: Preliminary results of social positioning and cognitive intervention. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.641765

Wellman, H. M. (2014). *Making minds: How theory of mind develops* (p. 366). Oxford University Press. viii.

Westby, C. E. (2005). Assessing and remediating text comprehension problems. In *Language and reading disabilities* (Second Edition,, *3* pp. 157–232). Pearson.

Westby, C. E. (2022). Play in children with Autism spectrum disorder. *Word of Mouth, 33*(5), 6–9. https://doi.org/10.1177/10483950221086338b

Whitely, C., & Colozzo, P. (2013). Who's who? Memory updating and character reference in children's narratives. *Journal of Speech, Language, and Hearing Research, 56*(5), 1625–1636. https://doi.org/10.1044/1092-4388(2013/12-0176)

Wolfer, P., Baumeister, F., Rudelli, N., Corrigan, G., Naigles, L. R., & Durrleman, S. (2024). Exploring metalinguistic awareness in school-aged autistic children: Insights from grammatical judgment. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-024-06569-y

Yau, S. H., Brock, J., & McArthur, G. (2016). The relationship between spoken language and speech and nonspeech processing in children with autism: A magnetic event-related field study. *Developmental Science, 19*(5), 834–852. https://doi.org/10.1111/desc.12328

Yu, B. (2013). Issues in bilingualism and heritage language maintenance: Perspectives of minority-language mothers of children with Autism Spectrum Disorders. *American Journal of Speech-Language Pathology, 22*(1), 10–24. https://doi.org/10.1044/1058-0360(2012/10-0078)

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research, 15*(5), 778–790. https://doi.org/10.1002/aur.2696