Research Article

Autism & Developmental Language Impairments

Autism & Developmental Language Impairments
Volume 8: 1–15

DOI: 10.1177/23969415231168557
journals.sagepub.com/home/dli

SAGE

Linguistic markers of autism spectrum conditions in narratives: A comprehensive analysis

Kristen Schroeder
Department of Philosophy, Classics, History of Art and Ideas (IFIKK), University of Oslo, Oslo, Norway
Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Catalunya, Spain

Joana Rosselló
Department of Catalan Philology and General Linguistics, Universitat de Barcelona, Barcelona, Catalunya, Spain

Teresa Ribalta Torrades
Fundació Orienta, Barcelona, Spain

Wolfram Hinzen
Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Catalunya, Spain
ICREA (Institute of Advanced Studies of Catalonia), Barcelona, Spain
Hermanas Hospitalarias, Benito Menni Hospital, Spain

Abstract

Background & Aims: Narratives are regularly elicited as part of standardized assessments for autism spectrum conditions (ASC) such as the ADOS, but have rarely been utilized as linguistic data in their own right. We here aimed for a specific and comprehensive quantitative linguistic profile of such narratives across nominal, verbal, and clausal domains of grammatical organization, and error patterns. **Methods:** We manually transcribed and annotated narratives elicited from the ADOS from a sample of bilingual autistic Spanish-Catalan children (n = 18), matched with typically developing controls (n = 18) on vocabulary-based verbal IQ. **Results:** Results revealed fewer relative clauses and more frequent errors in referential specificity and non-relational content-word choice in ASC. Frequent error types are also discussed qualitatively. **Conclusions & Implications:** These findings, based on more finegrained linguistically defined variables, help to disentangle previous inconsistencies in the literature, and to better situate language changes in the spectrum of neurocognitive changes in this population.

Keywords

ADOS, narrative, reference, language, grammar

Introduction

Although narrative speech data are freely generated as part of the Autism Diagnostic Observation Schedule (ADOS) resulting language samples have rarely been utilized as linguistic data in their own right. Narrative elicitation allows us to profile within the same task both formal aspects of grammar (such as tense and agreement) and how grammatical devices are used in referential strategies in the development of a story. In the

Corresponding author:
Kristen Schroeder, Department of Philosophy, Classics, History of Art and Ideas (IFIKK), Georg Morgenstiernes Hus, Office 504, Blindernveien 31, 0851 Oslo, Spain.
Email: k.f.schroeder@ifikk.uio.no

literature on language in ASC, analysis of formal syntax often relates to the discussion of specific difficulties with language in ASC (Boucher et al., 2012; Modyanova et al., 2017; Terzi et al., 2019) while differences in language use are discussed in relation to cognitive differences among this population, such as in mentalizing or theory of mind (ToM), cognitive flexibility (executive functioning) and global information processing (central coherence) (Happé & Frith, 2006). A priori, all of these cognitive mechanisms may relate to the organization of language at a discourse level: ToM, because it relates to our capacity to represent mental states, which enters in how we present information in discourse for an interlocutor; executive functioning, because it relates to managing information to be integrated over narrative; and central coherence, since information needs to be held together and be organized coherently around a topic. However, a comprehensive linguistic analysis of narrative data in ASC provides the essential linguistic baseline datapoints, which form the basis for testing specific predictions of neurocognitive models of ASC incorporating the foregoing constructs, and for assessing their explanatory scope (Hinzen, 2022).

Previous literature on narrative in ASC has revealed both variability and commonalities of ASC groups across studies relative to age-matched controls. For example, a reduction in syntactic complexity was found in some narrative studies (Banney et al., 2015; Norbury & Bishop, 2003), but not others (Colle et al., 2008; Diehl et al., 2006; Losh & Capps, 2003; Rumpf et al., 2012). There are indications of an increase in formal syntactic errors, yet also variability with regards to the way or extent to which such errors were assessed. For example, Norbury and Bishop (2003) found increased rates of tense-marking errors among their ASC group to be similar to controls with specific language impairment (SLI). In studies by Banney et al. (2015) or Rumpf et al. (2012), on the other hand, syntactic errors were either not under exploration as in the case of the latter or rather coarse-grained as in the former, such as assessing whether an error occurred at the lexical or utterance level. Looking towards commonalities, a meta-analysis by Baixauli et al. (2016) of studies of narrative in ASC found that measures of story cohesion were the most consistent area of weakness in the narrative of autistic children. How cohesive devices are defined varies across studies, though cohesion often broadly relates to the use of referential devices while constructing a narrative.

Aside from presenting more tense errors, the participants in Norbury and Bishop (2003) also differed from their typically developing peers on a standardized language assessment, the Clinical Evaluation of Language Fundamentals (CELF, Semel et al., 2003). This pattern of generalized weakness in productive and receptive language and presentation of SLI-like errors (i.e., morphological) has been characterized as autism with language impairment (ALI) (Kjelgaard & Tager-Flusberg, 2001), though the comparability of the language profiles seen in ASC and in SLI is also questioned (Boucher et al., 2012; Sukenik & Friedmann, 2018; Taylor et al., 2014). The other autistic groups in the studies surveyed here were similarly matched on chronological age but were not reported to differ from control groups on standardized verbal skills, however, comparability is limited as standard measures varied across studies. ALI entails a strict interpretation of structural language impairments, often at the sentence-level and which does not consider aspects such as cohesion. Narrative assessments have been demonstrated to uncover language difficulties that might not be noticeable through sentence-level tasks (King & Palikara, 2018) and therefore may help shed light on a more comprehensive profiling of language skills in autism more generally.

Against this background, our aim in the present study was to bring a more basic linguistic perspective to bear on the profiling of narrative in autistic children. We view this as prior to clarifying the ultimate cognitive basis of linguistic differences for this population and whether commonly posited neurocognitive differences in ASC in such constructs as ToM, weak central coherence, or executive functioning (Happé & Frith, 2006), are conceptually or empirically suited to explain them. A comprehensive and systematic linguistic analysis, carried out in terms of linguistically specific and independently motivated variables, could also provide a baseline also for assessing whether SLI is a suitable model for understanding language profiles in ASC. We, therefore, aimed to identify variables both within and across the main domains of structural organization of language (noun phrases, verb phrases, and clauses), along with error patterns.

In particular, in the case of clausal complexity, in previous studies, different clausal construction types have often entered indiscriminately in a diverse set of composite variables used to measure the construct of syntactic complexity, each of which included a range of different syntactic construction types with different functions (e.g., Norbury & Bishop, 2003; Norbury et al., 2014; Rumpf et al., 2012). This may have contributed to the aforementioned variability of previous findings. Cataloguing syntactic construction types by their functionality may prove a better strategy for grouping construction types in that their functionality may inform the cognitive processes involved. For example, in the syntactic domain of clauses, virtually any clause embedded as an argument specifies the content of a mental state or speech act (e.g., *The frogs thought they could fly*), constituting an immediate link between syntax and ToM that could test a ToM-based neurocognitive model of language functioning in ASC, as independently motivated by experimental evidence between argument clauses and ToM capacities (Polyanskaya et al., 2022; Schroeder et al., 2021). On the other hand, relative clauses (e.g., *The frogs that could fly*) are adjuncts grammatically, which add descriptive specificity to referential NPs, while (non-relative) clauses adjoined to superordinate clauses (e.g., *While they were flying, they saw a lady*; *They dropped down in order to rest*) relate events to each other, specifying e.g., temporal or causal relations between them. Targeting such distinct construction types in a more hypothesis-driven fashion is

therefore motivated for this population since it may help clarify differences in their language profiles.

The importance of referentiality for studies of cohesion similarly motivates distinguishing between different syntactic construction types within the domain of noun phrases (NPs). An increase in referential anomalies relative to TD controls or other clinical groups is among the most consistent findings among studies of narrative production in ASC – whether through an overproduction of ambiguous pronouns (Banney et al., 2015), a reduction of pronouns in favor of full definite noun phrases (NPs) (Rumpf et al., 2012), or an increase of both ambiguous nouns and pronouns (Norbury & Bishop, 2003). In the study by Banney et al. (2015), however, reference to characters was assessed through a measure of cohesion that tracked the means of introduction and maintenance of novel characters, through indefinites NPs, definite NPs, nouns only, pronouns, or omission. This measure did not distinguish between the ASC and TD groups.

As in the case of syntactic complexity, our aim in the domain of referentiality was not to start from a prior conceptualization of referential errors under the label of cohesive adequacy, but a comprehensive analysis of referential function in narrative across different NP types from a linguistic point of view. NPs systematically differ structurally, in line with how they function referentially in discourse (Hinzen & Sheehan, 2015). Thus, generic NPs are maximally nonspecific (e.g., *Frogs can fly*), conveying general knowledge only. Indefinite NPs can but need not feature determiners such as 'a' and introduce individuals into the discourse (e.g., *I see frogs; A lady sleeps*). Definite NPs must have determiners (articles, demonstratives) and identify specific objects (e.g., *Now the frogs fall to the ground, The frogs flew into these sheets*) or anaphorically refer to them, by resuming a referent previously introduced. As for pronouns, they can be deictic when they refer to the nonverbal, situational context (e.g., *The frogs flew into these, It/she/he is stuck*), or anaphoric, which is the predominant function for 3rd person pronouns. In this way, there is a whole system of NPs different in structure and function, which needs to be coordinated and together serves to set up and maintain a space of referents between the speaker and listener, based on shared lexical concepts (e.g., 'frog').

A linguistic analysis of narratives targeting this referential system as a whole has not yet been carried out to our knowledge. In experimental studies, however, definite NP comprehension weakness have been repeatedly noted (Fine et al., 1994; Modyanova, 2009; Schaeffer et al., 2014; Schroeder, 2019), and the classical case of pronoun reversals in ASC (Kanner, 1943) shows that problems with deictic forms of definite NPs are by no means restricted to problems of referential cohesion in narratives (see also Bartolucci et al., 1980; Mizuno et al., 2011; Shield et al., 2015). In fact, these may form just the tip of an iceberg, as referentiality is a dimension of language well-known to be affected throughout the spectrum, from the complete absence of declarative reference in non- or minimally verbal autistic people (Hinzen et al., 2020), to associative rather than referential forms of word use (Preissler & Carey, 2005; Preissler, 2008), anomalous lexical development (Arunachalam & Luyster, 2016), and delayed declarative pointing (which is closely related to language development: Colonnesi et al., 2010).

Moreover, referentiality in ASC has barely been explored in the verbal domain. In verb phrases (VP), the same events as depicted on the page of a story book can always be described using either 'static' verbs (e.g., *The frogs are up in the air*, *They wear capes*, *They feel happy*), or else more dynamic 'episodic' ones (e.g., *The frogs rise to the air*, *They are picking up sheets*). Aside from being more complex in terms of lexical and grammatical aspect, the choice of specific VP types could therefore be cognitively revealing: an event that is perceptually identical can still be conceptualized differently, as revealed through either static or dynamic verbalizations and the different forms of linguistic complexity involved in these. Conceptualizing visually presented events, from a certain level of complexity onwards, can be cognitively highly challenging for children of young chronological or verbal mental ages (Hinzen et al., 2022; see also de Villiers, 2014; Shukla & de Villiers, 2021). VP types were therefore included in our annotation scheme.

In sum, there is significant evidence that elicitation of spontaneous narratives is a sensitive tool to assess language anomalies across all broad domains of structural language in ASC. However, from a linguistic perspective, previous approaches have neither been very systematic (comprising construction types across nominal, verbal, and clausal domains), nor very specific (e.g., using generic or else composite measures, which conflate what often are linguistically very different phenomena, such as passives and embedded clauses). The fact that ASC can end up indistinguishable at a group level from SLI on language impairment (Norbury & Bishop, 2003; Norbury et al., 2014) motivates a deeper exploration using a fine-grained and comprehensive set of linguistic variables, which can provide potential differentiation.

Present study

We used ADOS-based assessments of narrative competence with the Tuesday book (Wiesner, 1991), both to allow for comparability with several previous studies and exploit linguistic data freely generated as part of clinical assessments. A new annotation scheme using linguistically defined variables was created to address the following specific research questions:

(i) Can groups be distinguished by the quantitative distribution of syntactic construction types in a-c?
   (a) NPs (pronouns/clitics, anaphoric and nonanaphoric definite NPs, and indefinite lexical NPs);
   (b) VPs (static vs. episodic);

(c) Clauses (embedded as adjuncts to NPs (relative clauses) and as clausal adjuncts or arguments (complement clauses)).

(ii) Can groups be distinguished by error patterns according to either a or b below?
   (a) Grammatical-level, further subdivided further into formal-syntactic errors and errors in the referential use of grammatical devices such as NPs;
   (b) Lexical-level errors, further subdivided into relational, e.g., *but*, vs. non-relational, e.g., *carpet*.

To explore the cognitive basis of the linguistic patterns in question, correlations with verbal IQ, as assessed by the Peabody Picture Vocabulary Test (PPVT-III, Dunn et al., 2010), were also run, as well as with IQ measures (IQ, working memory and processing speed) in the ASC group, as assessed by the WISC-IV (Wechsler, 2003). We predicted that despite individual matching of groups on a language criterion, linguistic vulnerability in the ASC group would show in:

(i) A quantitative reduction of embedded clauses (cf. Banney et al., 2015), specifically in relative and complement clauses (Durrleman & Delage, 2016; Durrleman et al., 2018), as well as of anaphoric NPs, including both third-person clitic pronouns and definite-anaphoric lexical NPs.
(ii) More grammatical-level errors, particularly referential ones relating to specificity as required in definite-anaphoric and definite-non-anaphoric NPs, pronouns, clitics, and verbal inflection, with a potentially spared lexicon (e.g., Boucher et al., 2012).

Methods

Participants

Eighteen children (mean age: 9;11, range: 7;4–12;6 years old) diagnosed with ASC participated in this study. ASC participants were required to score positively for autism in either the ADOS-G or the autism diagnosis interview-revised (ADI-R), as well as have a non-verbal IQ above the threshold for intellectual disability, as measured by the Wechsler Intelligence Scale for Children (WISC-IV). These assessments were conducted at the time of the data collection by the research team. All children were native speakers of Spanish and Catalan. ASC participants were recruited from a network of juvenile mental health centers in the Baix-Llobregat county of Barcelona, a region that tends to be Spanish-dominant. ASC participants were individually matched with 18 typically developing controls based on receptive vocabulary verbal IQ, as measured by the PPVT-III (see Krasileva et al., 2017, for an assessment of the validity of using the IQ generated by the PPVT as a proxy for verbal IQ for autistic individuals). Controls were identified as being a match if they were within one chronological year of age and five verbal IQ points. Controls were recruited from an elementary school in the Berguedà region of Barcelona. As the Berguedà region is Catalan-dominant, we further required participants to speak Spanish with at least one caregiver. All ASC participants spoke Spanish as a home language. Male to female sex ratios were similar across groups: the ASC group was comprised of 17 males and one female and the control group was comprised of fifteen males and three females.

Table 1. Participant information.

	ASC (n = 18) Mean (SD)	TD (n = 18) Mean (SD)	t	p
Sex ratio (male:female)	17:1	15:3	n/a	n/a
Verbal IQ (PPVT)	99.33 Range: 55–120	99.94 55–122	−.112	.911
Chronological Age	9;11 7;4–12;6	9;9 6;8–12;10	316	.754

Table 2. ADOS and WISC-IV means and SD for ASC participants (n = 18).

ADOS-G	Mean (SD)	ADI-R	Mean (SD)	WISC-IV	Mean (SD)
Communication Cutoff: 2	2.5 (1.73)	Communication Cutoff: 8	12.1 (4.8)	Full IQ	103.8 (22.1)
Social Interaction Cutoff: 4	4.9 (2.5)	Reciprocal social Interaction Cutoff: 10	13.9 (5.7)	Verbal comprehension	104.7 (20.3)
Communication + social interaction Cutoff: 7	7.5 (3.7)	Repetitive behaviors and restricted interests Cutoff: 3	4.3 (3)	Perceptual reasoning	105.8 (20.9)
Repetitive behaviors and restricted interests (No cutoff)	1.5 (1.4)	Evidence of symptoms before 36 months of age Cutoff: 1	2.7 (2)	Working memory	101.4 (20.5)
				Processing speed	97 (13.6)

Independent t-tests showed that there were no significant differences between groups on either verbal IQ or age (see **Table 1**). Profiling of ADOS-G and WISC scores for the ASC group are included in **Table 2**. This study was approved by the Vall D'Hebron University Hospital Ethical Board in Barcelona, Spain. Written informed consent was obtained from parents for all participants and additionally by adolescent participants themselves.

Narrative elicitation procedure

All participants were tested individually in a quiet room and video-recorded so that their narrative production could be later transcribed and annotated. The participants were presented with the picture book *Tuesday* (Wiesner, 1991). Following the ADOS-G (Lord et al., 2000) administration procedure, the examiner began narrating the first page and then allowed the child to continue the narration until the final pages when the experimenter would take over again. The experimenter provided neutral feedback. If a child didn't begin or stopped in the middle of the narrative, they were asked to continue the story. The book depicts a fanciful story in which frogs fly through a town at night on lily pads.

Coding scheme

Syntactic complexity in the nominal, verbal and clausal domains. The narratives were transcribed and annotated using CLAN (MacWhinney, 2000). All narratives were divided into utterances, which were defined as self-standing (grammatically independent) units of discourse providing new information. In the interest of profiling how the ASC participants refer to entities in the story and maintain reference, we categorized each NP according to NP type. NPs were first classified as either indefinite (e.g., *some frogs*) or definite (e.g., *the frogs*; *he*). Definite NPs were further classified as anaphoric (e.g., *they*; *the frogs she had seen*) or not (e.g., *this/a policeman*). Clitic pronouns were tagged as + clitic in addition to being classified as definite-anaphoric, with a further judgment of whether they were 'successful' or 'unsuccessful', depending on whether the clitic was involved in a referential error or not. Clitic errors were specified in order to be able to contrast the ratio of referential errors that were pronominal in nature. Finite VPs were classified as to whether they were stative or not (episodic). Syntactic complexity was scrutinized based on tagging clausal embedding and distinguishing between relative clauses, complement clauses, and (non-relative) adjunct clauses (see **Table 3**, below). Generic productive measures such as number of words as well as number of utterances were also computed, for comparability with previous studies (e.g., Banney et al., 2015 and Rumpf et al., 2012) and contrasts between generic and linguistically more specific measures.

As the length of narratives varied across participants, frequencies of all variables were normalized by the total number of words, with the exception of embedded clauses, which were normalized by the total number of utterances. We also compared the proportional use of each grammatical device type relative to its domain, creating nominal, verbal, and clausal domain ratios. For example, the definite-anaphoric NP ratio was calculated by dividing the total number of definite-anaphoric NPs by the total number of NPs.

Errors. We coded errors within the nominal (**Tables 4**) and verbal domains (**Table 5**). We also created composite score variables, specifically: total errors, total lexical errors, total grammatical errors, total verb errors, total nominal errors, and finally total nominal reference errors.

Mental state content. Mental state and emotion words were annotated as 'congruent', if used correctly or plausibly, or 'incongruent', if there was a clear mismatch between the emotion or mental state described by the participant and what was depicted in the illustration.

Reliability. Ten percent of the transcripts were randomly selected and transcribed again by a new transcriber; another ten percent of transcripts were independently annotated. Prior to reliability calculation, the first author met with the independent transcriber and annotator to establish a consensus of disagreements. The total number of point-by-point agreements were divided by the sum of the total agreements and disagreements, resulting in an inter-transcriber reliability of 90.3% and an inter-annotator reliability was 90.1%.

Table 3. Linguistic measures of construction types across syntactic subdomains.

Nominal Domain	Verbal domain	Clausal domain
Indefinite NP/words	Stative verb/words	Relative clauses/utterances
Definite NP/words	Episodic verb/words	Complement clause/utterances
Definite NP-non anaphoric		Adjunct clause/utterances
Definite NP-anaphoric		
Clitic-successful		
Clitic-unsuccessful		

Table 4. Nominal domain error variables.

Lexical level errors	
Wrong Lexical Item (non-relational)	(1) Imprecise content word, which may either be incorrect such as saying *parachute* rather than *cape*, or overly general such as saying *clothes* rather than *cape* (in the context in which the frogs wear a cape like that of superman). (2) neologism (e.g., *superdonde* for a frog in a cape, like a superhero) *y luego los sapos cogen una alfombra* [And then the frogs take a carpet] Comment: the frogs actually get caught in a bed sheet.
Wrong Lexical Item (relational)	Imprecise word choice, particularly in relation to functional words such as prepositions, conjunctions, etc. *Ladrarlos a perseguirlos* [Bark them to chase them (**a* rather than *para*)] Comment: in this case, the preposition *a* is incorrect as it does not impart intention. The correct preposition in this case would be *para*. We also note a secondary grammatical error, addressed in 'verb error'.
Grammatical level errors	
Vague Reference	Reference to entities that is too vague in nature to allow the listener to capture the scene well. *Aquí pasa algo* [Something is happening here].
Failed Specific Reference	Misuse of pronominal forms of reference (including person inflection), introducing a new referent with an unprecedented pronoun, or otherwise unclear who or what the intended referent is). *Ellas estaban viendo la tele y después se fue* [they were watching TV and then left-singular).] Comment: in the picture, one frog of the group left
Definite/Indefinite mismatch	Use of a definite or indefinite article in an inappropriate context. *Se quedaron horas y horas viendo la televisión. La rana la rana levitadora iba tranquilamente por el jardín* [[the frogs] stayed hours and hours watching television. *The frog the flying frog went slowly through the garden]
Noun agree error	Error in noun agreement (e.g., gender or plural). *Y por la mañanas cayeron al agua* [And in the$_{singular}$ morning$_{plural}$ they fell in the water]
Verb error	Error in verb agreement (e.g., person or number) or licensing errors such as making a non reflexive verb reflexive) *Ladrarlos [ladrar a ellos]* *Bark them [bark at them]

Table 5. Verb domain errors.

Verbal Domain Errors	Example
Lacking tense in the main clause	*Un señor cenando y mira los sapos pasar* A man eating and looks at the frogs passing
Wrong or inconsistent tense or aspect	*Cuando vuelan los sapos en las hojas encontraron un pequeño pueblito* When the frogs fly they found a tiny little town
Error in the verb root or conjugation	**ponieron (pusieron)'* *They putted (they put)
Truncation	*Se ha con la manta* She had with the blanket

Analysis plan. Statistical analyses were carried out using SPSS version 24. Where assumptions of normal distribution were met, independent t-tests were run to assess group differences. Where there were significant outliers or data were skewed as assessed by the Shapiro–Wilk test, we carried out non-parametric Mann–Whitney U-tests and compared medians. To adjust for multiple comparisons and strike an appropriate balance between a risk of type I and type II errors, we adjusted the alpha-level of significance by dividing .05 by the number of variables involved in a domain where comparisons were made: the clausal, nominal and verbal domains, and errors, leading to significance thresholds of .005 for both the grammatical profiling and error variables and a threshold of .01 for the remainder of variables. Only results meeting these significance thresholds, as well as statistical tendencies, are reported below. This quantitative analysis is followed by a qualitative analysis of two specific observed error patterns, which were added post-hoc to this analysis plan in order to provide further insight into a specific pattern that transpired during the annotation process.

Results

Quantitative analyses

Syntactic complexity. ASC participants produced significantly fewer embedded clauses: 42.9% of utterances in the TD group had an embedded clause in comparison to 16.8% in the ASC group. In particular, the ASC group produced significantly fewer relative clauses, while lower frequency of (non-relative) adjunct clauses was at the level of a tendency (see **Figure 1** and **Table 6**).

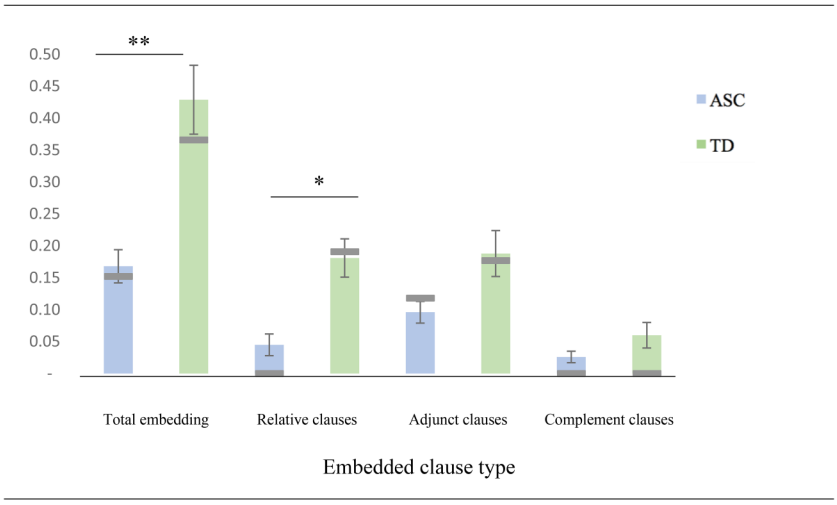

Figure 1. Rate of embedded clauses per utterance: means and medians (gray bars). *indicates $p = .001$, **$p = .0001$.

Table 6. Group comparisons of nominal, verbal and clausal domain variables using Mann-Whitney U-tests.

	ASC Median Mean (SD)	TD Median Mean (SD)	*U*	*z*	*p*	Effect size
Generic Measures						
Total words	185.5 172.7 (60)	172.5 187 (45)	174	.380	.719	.063
Total utterances	17.5 17.6 (5.7)	13 13 (2.5)	80.5	−2.591	.009*	.432
Nominal Domain						
Indefinite	.057 .06 (.023)	.058 .057 (.023)	145.0	-.538	.606	.09
Definite non-anaphoric	.044 .049 (.025)	.047 .05 (.017)	170.0	.253	.815	.042
Definite anaphoric	.130 .136 (.027)	.137 .142 (.028)	181.5	.617	.542	.103
Total clitics	.011 .123 (.01)	.017 .015 (.01)	196	1.077	.293	.048
Clitic-successful	.009 .01 (.008)	.015 .014 (.01)	197.5	1.126	.265	.193
Clitic-unsuccessful	0 .002 (.004)	0 .001 (.004)	145	-.828	.606	.138
Verbal Domain						
Stative verbs	.038 .04 (.017)	.039 .039 (.018)	168	.19	.864	.032
Episodic verbs	.108 .113 (.023)	.124 .119 (.026)	182.5	.649	.521	.108
Clausal Domain						
Relative Clauses	0 .045 (.072)	.191 .181 (.128)	266.5	3.387	.001*	.565
Complement Clauses	0 .026 (.036)	0 .06 (.085)	191	1.025	.372	.062
Adjunct Clauses	.118 .061 (.073)	.177 .188 (.154)	224.5	1.990	.047	.332
Total Embedding	.152 .168 (.111)	.366 .429 (.227)	282.5	3.814	.0001**	.636

Table 7. Error scores normalized by number of words.

	ASC Median (range) Mean (SD)	TD Median (range) Mean (SD)	U	z	p	Effect size
Total errors	.038 (.013-.089) .041 (.022)	.015 (0–.03) .015 (.011)	40	−3.860	.0001***	.643
Total wrong lexical item (WLI)	.012 (0–.04) .016 (.012)	.005 (0–.01) .004 (.004)	59	−3.284	.001**	.547
WLI relational	.003 (0–.018) .006 (.007)	0 (0–.01) .001 (.002)	101	−2.244	.055	.374
WLI non-relational	.011 (0–.031) .016 (.009)	.002 (0–.01) .003 (.003)	73	−2.884	.004*	.481
Total grammatical errors	.022 (0–.053) .025 (.015)	.009 (0–.03) .011 (.009)	70	−2.914	.003*	.486
Total verbal domain errors	.005 (0–.027) .007 (.008)	.002 (0–.02) .005 (.006)	142.5	−.646	.542	.108
Lacking tense in the main clause	0 (0–.01) .001 (.002)	0 (0–.01) .0 (.002)	161.5	−.040	.988	.007
Wrong or inconsistent tense or aspect	0 (0–.026) .004 (.007)	0 (0–.02) .003 (.006)	167	.184	.888	.031
Error in the verb root or conjugation	0 (0–.013) .001 (.004)	0 (.01) .001 (.003)	179	.828	.606	.138
Truncation	0 (0–.008) .001 (.003)	0 (0–0) 0 (0)	126	−2.087	.265	.348
Total NP errors	.017 (0–.047) .018 (.11)	.005 (0–.02) .006 (.006)	58	−3.298	.001**	.55
Noun agreement	0 (0–.017) .003 (.005)	0 (0–.01) .001 (.002)	117	−1.745	.161	.291
Verb agreement	0 (0–.009) .001 (.003)	0 (0–.01) .001 (.002)	162	0	1	0
Total referential errors	.011 (0–.037) .013 (.018)	.004 (0–.02) .004 (.005)	75	−2.792	.005*	.168
Vague Reference	0 (0–.018) .002 (.005)	0 (0–.01) .002 (.004)	180	.698	.584	.116
Failed Specific Reference	.007 (0–.037) .009 (.01)	0 (0–.01) .001 (.003)	71	−3.117	.003*	.52
Definite/Indefinite mismatch	0 (0–.017) .001 (.004)	0 (0–.01) 0 (.001)	144.5	−1.015	.584	.061

*indicates $p \leq .005$; *$p \leq .001$; **$p \leq .0001$.

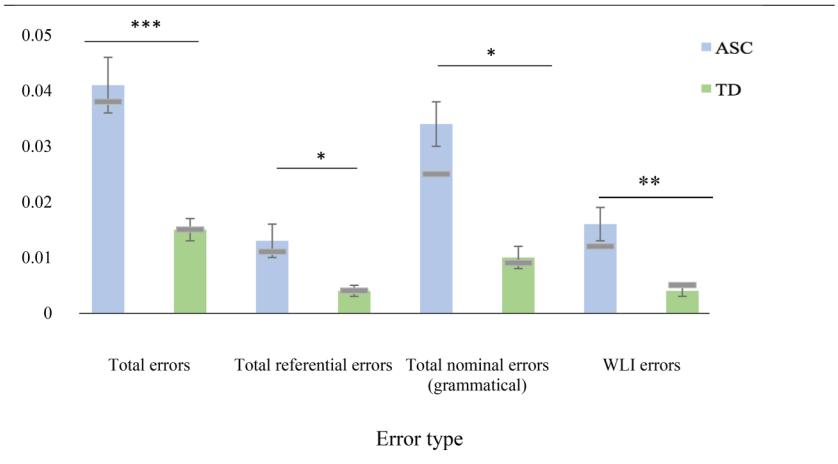

Figure 2. Frequency of error types: means and medians (gray bars). *indicates $p \leq .005$; *$p \leq .001$; **$p \leq .0001$. Gray bars indicate medians.

Errors. ASC participants produced lexical-level and grammatical-level errors at a significantly higher rate than TD controls (see **Table 7** and **Figure 2;** recall samples of errors in **Table 4**). Within grammatical-level errors, the ASC group produced significantly more total NP errors, specifically errors in the subcategory failed specific reference, which captured referential errors in which it is unclear who or what the intended referent is. Within lexical-level errors, ASC participants produced significantly more non-relational lexical errors (e.g., using an erroneous noun) in comparison to TDs. While the ASC group also produced relational lexical errors (e.g., using erroneous preposition) at a higher mean frequency than the TDs, this difference was only a tendency ($U = 59$; $z = -3.86$; $p = .055$).

Results summary. Several significant differences across groups emerged, including (1) significantly more utterances in the ASC group while having similar quantities of total words; (2) a reduction in the rate of embedded clauses in ASC in relation to the TD, particularly in relative clauses but not (non-relative) adjunct clauses or complement clauses; (3) a greater rate of referential errors in ASC, particularly that of failed specific reference; (4) a greater rate of lexical errors in ASC, particularly in non-relational lexical choices. In contrast, the ASC participants did not differ significantly with regard to either the frequency of different NP and VP-types. Nor did they differ in terms of the frequency of verbal errors, or the rate of use of internal state language across groups. Regarding correlations, no linear relationships were identified between verbal IQ and key linguistic variables significant in the group comparisons, specifically failed specific reference, wrong lexical items (WLI), and embedded clause frequency. In the ASC group, none of the cognitive measures available as part of the WISC, including working memory and processing speed related to the above key linguistic variables (see **Table 8**).

Qualitative analyses

We provide a detailed qualitative analysis of two specific patterns of failed specific reference based on four examples. Consider (1) and (2), where a star in the translation localizes the error occurring—in bold in the original utterance and gloss:

(1) *…se fueron saltando al lago y se tiraron de cabeza y están allí, como pensando,* ***está enfadado****.*

Went$_{3p-pl}$ jumping back to the lake and jumped$_{3p-pl}$ head-first and are$_{3p-pl}$ there, like thinking, **is$_{3p-sing}$ angry.**

'(They) went jumping back to the lake and they jumped head-first and they are there, like thinking, (*he) is angry.'

(2) *Aquí* ***vuela*** *tranquilo y aquí se choca con un perro aquí se xxx [unintelligible] con un perro y* ***se van*** *y se* ***los come.***

Here calmly **flys$_{3p-sing}$** and here runs into a dog here xxx with a dog and **go$_{3p-pl}$** and **clitic:3p-pl**. eats$_{3p-sing.}$

'Here (he) calmly flies and runs into a dog here xxx with a dog and they go and he eats *them.'

In these examples, the children appear to be referring through inflections and clitics to a specific individual or group, yet get entangled in how person and number in inflection and clitics are to be used in conjunction over a single stretch of discourse, so as to convey a coherent content. The pictures aid the listener in forming hypotheses about what is intended, and the use of deictics such as *aquí* and *allí* (here and there) indicates reliance on the shared visual context for referencing. Reference appears to be 'exophoric' in this sense, i.e., reliant on an entity in the shared visual domain, rather than 'endophoric', i.e., picking up a referent introduced previously into the discourse—it is as if the exophoric scope is continually being refreshed as the utterance proceeds.

In another error pattern, the participant's utterance in its literal interpretation does not match a viable reality in consideration of the visual support provided. Consider (3) and (4):

(3) *Entraban$_i$ por la ventana y* ***les$_i$ quitaban*** *los pañuelos.*

They$_i$ entered through the window and **(3rdPplural-CL$_i$) they$_a$** took the handkerchief*s*.

They$_i$ entered through the window and **they(-others) took** the handkerchiefs (from them$_i$)

(4) *Cogieron el mando sin que se* ***den*** *cuenta y estaban viendo la tele.*

They$_i$ took the remote control without noticing$_{i^{-}3p-pl}$ and were watching TV.

They took the remote control without noticing and were watching TV.

Table 8. Correlation analyses, Spearman's rho.

	Total Embedding	Total Errors	Wrong lexical Item	Failed Specific Reference
VIQ (PPVT-III)	r = −08, p = .754	r = −.362, p = .140	r = −.082, p = .747	r = −.205, p = .415
IQ (WISC-IV)	r = .210, p = .419	r = −.217, p = .402	r = −.034, p = .896	r = −.217, p = .402
Perceptual Reasoning	r = .268, p = .282	r = −.119, p = .638	r = .113, p = .598	r = −.163, p = .518
Working memory	r = .177, p = .483	r = −.278, p = .263	r = −.01, p = .969	r = −.181, p = .473
Verbal Comprehension	r = .129, p = .609	r = −.250, p = .316	r = −.168, p = .505	r = −.211, p = .4
Processing speed	r = −.063, p = .805	r = −.283, p = .805	r = .075, p = .767	r = −.36, p = .888

Here, there is a mismatch of what is being said with what is seen in the picture. In (3) we see the frogs entering through the window no longer wearing capes. The child is saying that they entered the house through the window and that *they* took the capes off them. But it is unclear who *they* would be. More likely, the intended meaning was to say *se quitaron los pañuelos* (they took off their capes), a clitic-V-NP possessive construction where the clitic refers to the possessor of the entity denoted by the NP, which then would be appropriately co-referential with the subject of the coordinate clause and, in turn, with that of the initial clause. In (4), the illustration presents an older woman sleeping in front of her TV while several frogs are watching TV and one is controlling the remote with his tongue. In the utterance, there is a number mismatch, however, as the verb *to notice* should be singular rather than plural, if it is to refer to the woman. Instead of reference to the frogs and the woman to be disambiguated, we find the plural number inflection uniformly across the utterance. There also is a tense/aspect error as the verb *den cuenta* is in the present tense, which is incongruent with the past tense just prior in the utterance. Errors of the type seen in (3) and (4) suggest difficulty in handling and coordinating linguistic devices used in order to refer to entities at a grammatical level.

To further explore this distinction between errors that may relate to maintaining an exophoric scope and those which suggest challenges in the grammatical build-up of the child's discourse, we further categorized and counted all instances of failed specific reference across both groups. Total instances and proportions are presented in **Table 9** below.

Discussion

These results reveal specific linguistic differences in the narratives produced by autistic children in comparison to TD peers of a similar age and matched on vocabulary-based verbal IQ. In this section we will discuss the syntactic complexity results first, followed by lexical-level and grammatical-level errors.

Syntactic complexity

In line with predictions, the ASC group produced fewer embedded clauses, and particularly fewer relative clauses. At the level of a tendency, there was a reduction in (non-relative) adjunct clauses as well. Complement clauses were rarely used in either group, so strength of comparisons between groups may have been limited. This result highlights the significance of clausal embedding as such for distinguishing these groups. The pronounced reduction in relative clauses in turn adds to a pattern of differences in the nominal domain as identified in the previous literature (Banney et al., 2015; Norbury & Bishop, 2003). Relatives embed under nouns and function so as to provide further descriptions of the referent of the NP. A reduction in relative and (at the level of a tendency) clausal adjuncts may thus form part of a pattern of reduced referential specificity (see more below), at the level of the object- and event-reference, respectively.

On the other hand, and contrary to our predictions, there were no group differences in the quantitative distribution of either NP or VP types as distinguished here. We speculate that this may relate to the use of a picture-book task, where the referents are visually present as and when they were being referenced. This allows definite NPs (whether full lexical NPs, clitics or null/inflectional) to be used exophorically, i.e., to refer to an entity in the shared visual domain, rather than endophorically, i.e., picking up a referent introduced previously into the discourse. While we distinguished definite-anaphoric NPs from definite-non anaphoric NPs (where a novel referent is introduced with a definite article), our annotation scheme therefore cannot fully exclude that some of the former (anaphoric) cases were in fact exophoric uses disguised as anaphoric NPs. This hypothesis is supported by previous findings from experimental studies that it is precisely (truly anaphoric) definite NPs that cause problems in ASC populations (Modyanova, 2009; Schaeffer et al., 2014; Schroeder, 2019). This predicts that they might be under- or mis-used in narrative production as well. Future results from studies of narratives both with and without visual support can address this crucial open question.

Table 9. Failed specific reference sub-type proportional use per group.

	ASC		TD	
	count	percent	count	percent
Exophoric scope	20	69	5	83.3
Grammatical error	8	27.6	1	16.7
Other	1	3.5	n/a	n/a
Total	**29**		**6**	

Error patterns

In line with predictions, error patterns distinguished groups, yet contrary to predictions, the lexical level (word choice) was not spared. A higher rate of lexical errors in the ASC group specifically appeared in what we called 'non-relational' lexical errors, which were virtually non-existent in the TD group. This was particularly surprising, as we matched participant groups on receptive vocabulary level. The finding also contrasts to claims in the literature that vocabulary is relatively preserved compared to grammar among autistic children without intellectual disability (e.g., Boucher et al., 2012; Eigsti et al., 2007; though see Arunachalam & Luyster, 2016). Wrong lexical choices (WLI) were often related to the target words in terms of appearance or function yet not quite matching the visual

stimuli or the previous representation of the object (e.g., naming capes or sheets worn by the frogs as *handkerchiefs* or *napkins*, as validated by the immediate visual context, but not the discourse and story (this type of error further adds to our above suggestion that the ASC participants were referring exophorically, not relating the current scene with what had passed in their own discourse previously). Other common lexical errors included mismatches between the lexical choice and the visual stimuli, such as referring to the hang-drying sheets as *blankets*. Given that both groups were highly comparable on receptive vocabulary as measured by single-word comprehension (i.e., the PPVT-III), we speculate that anomalies of lexical choice likely relate to the constraints of lexical retrieval within grammatical structures as built online over narrative time. While cognitive capacities in ToM are often drawn upon to provide a potential explanation for referential errors in ASC narratives (e.g., Colle et al., 2008; Suh et al., 2014; though see Arnold et al., 2009 for a contrary view), it is unclear how mentalizing difficulties would account for lexical errors as found here. In particular, identifying accurately a picture in the story such as a *sheet* or a *blanket* would not seem to depend on the mental state of the listener of the story.

The role of failed specific reference

Among grammatical-level errors, it is noteworthy that morphosyntactic errors did not distinguish groups and were indeed rare. On the other hand, failed specific reference occurred significantly more frequently in the ASC group. Unlike the lexical errors above, these errors do not relate to the *kind* of object being referred to (i.e., how it is conceptualized or described lexically), but rather to mishandling the means through which it is being referenced, resulting in a referential breakdown. As described in the introduction, NP types in human language form a system, which provides means to create and maintain a shared system of reference over narrative time. Grammatically distinct NP types support distinct forms of reference, and system-internal constraints guide their co-occurrence and relative ordering. For example, pronouns lack lexical descriptive content and depend for their appropriate use on either a prior introduction of the referent through a lexical (definite or indefinite NP), or else a perceptually salient referent in the narrative content (the exophoric case). Previously documented difficulties in handling such devices in ASC (such as under-use of pronouns, Arnold et al., 2009) could indicate a breakdown of this coordinated system, but it could also be due to system-external deficits such as deficits in ToM (e.g., Colle et al., 2008).

Against the latter option, Arnold et al. (2009) already argued that coordinating the referential system serves to structure both the narrator's own representation of the discourse and to guide that of the interlocutor. Investigating this issue further in Spanish is useful, as strong (i.e., independent, non-clitic) pronouns in this language are actually mainly used for emphasis and are not the default option, which instead is the use of verbal morphology (inflection) and clitics for subjects and objects (direct and indirect), respectively. Both inflection and clitics encode person and number features (and structural case features in clitics), which enable reference in pronouns and, in this regard, they are as pronominal as strong pronouns. In this respect, we found that both groups had a similar frequency of clitics, both quantitatively and in terms of errors in them, along with similar frequencies in lexical NP types as discussed above. Overall, this may suggest that failure of specific reference in our study is best interpreted as reflecting a weakness in coordinating the referential system of grammar as a whole over narrative time. Insofar as this system is subject to intricate and system-internal constraints on how various grammatically different referential devices are to be handled together, this conclusion illustrates a challenge for current neurocognitive models of ASC and for understanding language impairment in ASC. The absence of morphosyntactic errors together with problems of referential specificity do not seem typical of an SLI-type language impairment. In turn, while difficulties in the domain of ToM might be thought to specifically explain difficulties with pronouns (Colle et al., 2008), this is less clear for the pattern above.

Our qualitative analysis of error patterns in failed specific reference further contributes to the foregoing interpretation, as it suggests an increased reliance on an exophoric narrative style where reference is implemented via visually provided referents, and processing difficulties specifically appear in the coordination of referential devices in complex sentences. Errors such as those described in examples (3) and (4) (Section Qualitative analyses) have the effect of confusing the listener, hampering the build-up of a coherent discourse representation. The cause of the problem appears to be a difficulty in the speaker's handling of a system of referential devices and the constraints on using inflections and clitics in the context of other NPs over the course of a complex utterance. The overall rate and proportion of these errors (see **Table 9**) suggests that those related to an exophoric scope were the most prevalent. Non exophoric errors were also prevalent in the ASC group, though to a lesser degree, and rare in the TD group, with only one instance. While it is important not to over-interpret these findings from a qualitative assessment of error patterns, they further weaken an explanation of error patterns based on the notion of deficits in monitoring an interlocutor's mental state. Moreover, mentalizing difficulties in the ASC group with regard to internal state vocabulary were not observed. This is not to preclude that our participants may present mentalizing deficits in other tests or daily life. There is evidence that clausal embedding (including in relative clauses) may play a causal role in mentalizing and

meta representational capacities more generally (Durrleman et al, 2018; Schroeder et al., 2021). Reduced levels of embedded clauses in narrative production may indeed relate to increased mentalizing difficulties.

Cognitive capacities such as verbal working memory, too, may be relevant to explore in relation to referential errors, as they have been argued to relate to the ability to maintain discourse representations and anaphoric links in ASC (Schaeffer, 2018). Interestingly, Stanford et al. (2019) show that training of working memory capacities can lead to gains in syntax and specifically in clitic placement, in language-impaired groups. Note, on the other hand, that a deficit in working memory would not naturally incorporate the lexical errors that we found. Furthermore, for the ASC group, cognitive measures were available as part of the WISC, including working memory and processing speed. None of these cognitive variables related to these key linguistic variables, suggesting some autonomy of the latter, at least in this part of the spectrum (without intellectual disability).

A number of non-significant group differences in our study contribute to a comprehensive linguistic profile of this group as well as highlighting specific strengths in their narrative development. We explored the verbal domain, where neither the distribution between episodic and stative VPs nor error patterns distinguished groups. Here we found that ASC participants presented episodic descriptions of the scenes with well-developed morphosyntactic skills. This highlights the importance of the *nominal* domain when building a linguistic profile potentially specific to ASC, since both relative clauses and referential errors relate to this domain. Comparing the strengths and weaknesses found here in narratives by autistic children to children with developmental language disorders may further refine our understanding of ASC linguistic profiles.

Overall, this analysis of narratives in diagnostically relevant ASC assessments has shed light on a distinctive linguistic profile comprising differences in the quantitative distribution and qualitative use of specific syntactic construction types, and errors in both lexical choices and in managing a system of referential devices involved in generating referential specificity over narrative time. This profile imposes a linguistic constraint on current and future cognitive models of ASC. In particular, it is difficult to see it naturally explained by mentalizing difficulties. Indeed, future neurocognitive models of ASC need to be sensitive to the possibility that language impairment in ASC may not be secondary to a set of other cognitive impairments viewed as primary, nor be 'specific' in the sense of reflecting SLI as traditionally conceived. Independent evidence of divergences between the ASC and SLI linguistic profiles, consistent with evidence from the present study, further supports the latter possibility. Instead, it may be that, in the context of ASC, language itself as an essential social-communicative tool mediates certain social-cognitive abilities, and is of relevance in neurocognitive models of ASC for this reason (Hinzen, 2022; Hinzen et al., 2020). Recent evidence that training in the domain of embedded clauses can improve ToM capacities (Durrleman et al., 2019) is a further pointer in this direction: language and cognitive domains traditionally viewed as non-linguistic may be closely linked, and potentially share a cognitive substrate (Schroeder et al., 2021). The same may apply to referential functions. Dependent on both lexical choices and handling these choices within a complex interdependent system of grammatical devices, reference exploits virtually the entire complexity of the linguistic system. Reference is vulnerable across mental health conditions (Hinzen, 2017), and whether impairment in non-linguistic cognitive domains can explain this breakdown, or else language itself is the crucial explanatory factor, is an important open question.

Summary

Linguistic analysis of spontaneous narratives regularly elicited as part of standardized assessments for ASC shows that bilingual autistic Spanish-Catalan children without intellectual disability diverge from their neurotypical verbal-IQ-matched controls on specific linguistic variables: fewer adjunct clauses and more frequent errors in referential specificity and non-relational content word choice, jointly with an absence of morphosyntactic errors. These findings invite new inquiries on how to situate language changes within neurocognitive models of ASC, as they appear to define a distinctive linguistic profile, which resists explanation through familiar nonverbal cognitive impairments in ASC as much as assimilation to a comorbid SLI.

Limitations and future work

Inclusion of a comprehensive language test would have allowed us to assess how our linguistic measures relate to linguistic measures used to identify a typical SLI profile as well as determine whether linguistic differences identified would have transpired against an even tighter linguistic group matching. In order to further explore the effects of online language processing differences, further studies may benefit from profiling groups on both expressive language skills as well as receptive vocabulary. Another limitation is that only a part of the autism spectrum was targeted here, given our IQ inclusion threshold, which shields an extensive range of language impairment in ASC from view (Hinzen et al., 2020; Jack & Pelphrey, 2017). Third, participants of this study were all bilingual, which complicates comparability with previous studies of monolinguals. While new evidence supports that bilingualism can enhance cognition in ASC without intellectual disability (Durrleman et al., 2022), it is an important question whether it also impacts on linguistic variables extracted here. Future work should also consider the practical implications of the present study. Many of the linguistic differences

uncovered here are subtle (such as rate of embedded clauses), and not of a kind that necessarily meet a clinician's eye. Given their potential significance for training studies and diagnosis, objective assessments of language based on automated language processing options are therefore highly desirable. The measures used here do not yet lend themselves to such analyses.

Finally, future work should explore narratives or stories told by autistic children *without* visual support, as these may present greater interpretative challenges for the listener to overcome. As mentioned in the introduction, one of the aspects for which children are given narrative generation tasks is to assess sense-making skills. In order to build a narrative and draw inferences of how the story is unfolding, it is favorable to build endophoric representations which are continuous and involve anaphoric linking. If autistic children tend to favor exophora, as we suspect, we would expect non-visually supported narratives to show the characteristics identified here to a greater extent.

Credit author contribution statement

Kristen Schroeder: Conceptualization, Investigation, Formal Analysis, Writing- Original Draft, Writing - Review & Editing; **Joana Rossello**: Conceptualization, Writing- Original Draft, Writing - Review & Editing, Supervision; **Teresa Ribalta Torrades:** Resources; **Wolfram Hinzen:** Conceptualization, Writing- Original Draft, Writing - Review & Editing supervision, funding acquisition

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded through the Spanish Ministry of Economy and Competitiveness [FFI2013-40526P] and the grant PID2019-105241GB-I00/AEI/10.13039/501100011033 provided by the Ministry of Science, Innovation and Universities (MCIU) and the Agencia Estatal de Investigación (AEI) of Spain.

ORCID iD


Wolfram Hinzen https://orcid.org/0000-0002-6159-4215

References

Arnold, J. E., Bennetto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition*, *110*(2), 131–146. https://doi.org/10.1016/j.cognition.2008.10.016

Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research*, *9*(8), 810–828. https://doi.org/10.1002/aur.1590

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

Banney, R. M., Harper-Hill, K., & Arnott, W. L. (2015). The autism diagnostic observation schedule and narrative assessment: Evidence for specific narrative impairments in autism spectrum disorders. *International Journal of Speech-Language Pathology*, *17*(2), 159–171. https://doi.org/10.3109/17549507.2014.977348

Bartolucci, G., Pierce, S., & Streiner, D. (1980). Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. *Journal of Autism and Developmental Disorders*, *10*(I), 39–50. https://doi.org/10.1007/BF02408431

Boucher, J., Mayes, A., & Bigham, S. (2012). Memory in autistic spectrum disorder. *Psychological Bulletin*, *138*(3), 458–496. https://doi.org/10.1037/a0026869

Colle, L., Baron-Cohen, S., Wheelwright, S., & van der Lely, H. K. (2008). Narrative discourse in adults with high-functioning autism or Asperger syndrome. *Journal of Autism and Developmental Disorders*, *38*(1), 28–40. https://doi.org/10.1007/s10803-007-0357-5

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, *30*(4), 352–366. https://doi.org/10.1016/j.dr.2010.10.001

de Villiers, J. (2014). What kind of concepts need language? *Lang. Sci*, *46*(Part B), 100–114. https://doi.org/10.1016/j.langsci.2014.06.009

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

Dunn, L. M., Dunn, L. M., & Arribas, D. (2010). *PPVT-III – Peabody, Test de vocabulario en imágenes*. TEA Ediciones.

Durrleman, S., Burnel, M., De Villiers, J., Gibson, T. E., Yan, R., & Delage, H. (2019). The impact of grammar on mentalizing: A training study including children with autism spectrum disorder and developmental language disorder. *Frontiers in Psychology*, *10*, 2478. https://doi.org/10.3389/fpsyg.2019.02478

Durrleman, S., & Delage, H. (2016). Autism spectrum disorder and specific language impairment: Overlaps in syntactic profiles. *Language Acquisition*, *23*(4), 361–386. https://doi.org/10.1080/10489223.2016.1179741

Durrleman, S., Hinzen, W., & Franck, J. (2018). False belief and relative clauses in autism spectrum disorders. *Journal of Communication Disorders*, *74*, 35–44. https://doi.org/10.1016/j.jcomdis.2018.04.001

Durrleman, S., Peristeri, E., & Tsimpli, I. M. (2022). The language-communication divide: Evidence from bilingual children with atypical development. *Evolutionary Linguistic Theory*, *4*(1), 5–51. in press. https://doi.org/10.1075/elt.00037.dur

Eigsti, I. M., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders*, *37*(6), 1007–1023. https://doi.org/10.1007/s10803-006-0239-2

Fine, J., Bartolucci, G., Szatmari, P., & Ginsberg, G. (1994). Cohesive discourse in pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 315–329. https://doi.org/10.1007/BF02172230

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Hinzen, W. (2017). Reference across pathologies: A new linguistic lens on disorders of thought. Target article with peer commentary. *Theoretical Linguistics*, *43*(3–4), 169–232. https://doi.org/10.1515/tl-2017-0013

Hinzen, W. (2022). Rethinking the role of language in autism. *Evolutionary Linguistic Theory*, *4*(1), 129–151. https://doi.org/10.1075/elt.00040.hin

Hinzen, W., Peinado, E., Perry, S. J., Schroeder, K., & Lombardo, M. (2022). Language level predicts perceptual categorization of complex reversible events in children. *Heliyon*, *8*(7), e09933. https://doi.org/10.1016/j.heliyon.2022.e09933

Hinzen, W., & Sheehan, M. (2015). *The philosophy of universal grammar*. Oxford University Press.

Hinzen, W., Slušná, D., Schroeder, K., Sevilla, G., & Vila Borrellas, E. (2020). Mind–language = ? The significance of non-verbal autism. *Mind & Language*, *35*(4), 514–538. https://doi.org/10.1111/mila.12257

Jack, A., & Pelphrey, K. A. (2017). Annual research review: Understudied populations within the autism spectrum–current trends and future directions in neuroimaging research. *Journal of Child Psychology and Psychiatry*, *58*(4), 411–435. https://doi.org/10.1111/jcpp.12687

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*(3), 217–250.

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

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, *16*(2-3), 287–308. https://doi.org/10.1080/01690960042000058

Krasileva, K. E., Sanders, S. J., & Bal, V. H. (2017). Peabody picture vocabulary test: Proxy for verbal IQ in genetic studies of autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(4), 1073–1085. https://doi.org/10.1007/s10803-017-3030-7

Lord, C., Risi, S., Lambrecht, L., Cook, E.Jr, Leventhal, B., DiLavore, P., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

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

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd Edition). Lawrence Erlbaum Associates.

Mizuno, A., Liu, Y., Williams, D. L., Keller, T. A., Minshew, N. J., & Just, M. A. (2011). The neural basis of deictic shifting in linguistic perspective-taking in high-functioning autism. *Brain*, *134*(8), 2422–2435. https://doi.org/10.1093/brain/awr151

Modyanova, N. (2009). *Semantic and pragmatic language development in typical acquisition, autism spectrum disorders, and williams syndrome with reference to developmental neurogenetics of the latter*. Massachusetts Institute of Technology PhD thesis.

Modyanova, N., Perovic, A., & Wexler, K. (2017). Grammar is differentially impaired in subgroups of autism Spectrum disorders: Evidence from an investigation of tense marking and morphosyntax. *Frontiers in Psychology*, *8*, 320. https://doi.org/10.3389/fpsyg.2017.00320

Norbury, C. F., & Bishop, D. V. (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders*, *38*(3), 287–313. https://doi.org/10.1080/136820310000108133

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

Polyanskaya, I., Eigsti, I. M., Brauner, T., & Blackburn, P. (2022). Second-order false beliefs and linguistic recursion in autism Spectrum disorder. *Journal of Autism and Developmental Disorders*, *52*(9), 3991–4006. https://doi.org/10.1007/s10803-021-05277-1

Preissler, M. A. (2008). Associative learning of pictures and words by low-functioning children with autism. *Autism*, *12*(3), 231–248. https://doi.org/10.1177/1362361307088753

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, *97*(1). https://doi.org/10.1016/j.cognition.2005.01.008

Rumpf, A. L., Kamp-Becker, I., Becker, K., & Kauschke, C. (2012). Narrative competence and internal state language of children with Asperger syndrome and ADHD. *Research in Developmental Disabilities*, *33*(5), 1395–1407. https://doi.org/10.1016/j.ridd.2012.03.007

Schaeffer, J. (2018). Linguistic and cognitive abilities in children with specific language impairment as compared to children with high-functioning autism. *Language Acquisition*, *25*(1), 5–23. https://doi.org/10.1080/10489223.2016.1188928

Schaeffer, J., van Witteloostuijn, M., & de Haan, D. (2014). Article choice in children with high functioning autism (HFA) and in children with specific language impairment (SLI). *Linguistics in the Netherlands*, *31*(1), 107–128. https://doi.org/10.1075/avt.31.09sch

Schroeder, K. (2019). *Comprehension and production of referential expressions across Autism Spectrum Conditions*. (Doctoral dissertation, Universitat de Barcelona, Spain).

Schroeder, K., Durrleman, S., Çokal, D., Delgado, A. S., Marin, A. M., & Hinzen, W. (2021). Relations between intensionality,

theory of mind and complex syntax in autism spectrum conditions and typical development. *Cognitive Development*, *59*, 101071. https://doi.org/10.1016/j.cogdev.2021.101071

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals* (4th ed.). Psychological Corporation.

Shield, A., Meier, R. P., & Tager-Flusberg, H. (2015). The use of sign language pronouns by native-signing children with autism. *Journal of Autism and Developmental Disorders*, *45*(7), 2128–2145. https://doi.org/10.1007/s10803-015-2377-x

Shukla, M., & de Villiers, J. (2021). The role of language in building abstract, generalized conceptual representations of one- and two-place predicates: A comparison between adults and infants. *Cognition*, *213*, 104705. https://doi.org/10.1016/j.cognition.2021.104705

Stanford, E., Durrleman, S., & Delage, H. (2019). The effect of working memory training on a clinical marker of French-speaking children with developmental language disorder. *American Journal of Speech-Language Pathology*, *16*(4), 1–23. https://doi.org/10.1044/2019_AJSLP-18- 0238

Suh, J., Eigsti, I. M., Naigles, L., Barton, M., Kelley, E., & Fein, D. (2014). Narrative performance of optimal outcome children and adolescents with a history of an autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders*, *44*(7), 1681–1694. https://doi.org/10.1007/s10803-014-2042-9

Sukenik, N., & Friedmann, N. (2018). ASD Is not DLI: Individuals with autism and individuals with syntactic DLI show similar performance level in syntactic tasks, but different error patterns. *Frontiers in Psychology*, *9*, 279. https://doi.org/10.3389/fpsyg.2018.00279

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. (2014). Evidence for distinct cognitive profiles in autism spectrum disorders and specific language impairment. *J Autism Dev Disord.*, *44*(1), 19–30. https://doi.org/10.1007/s10803-013-1847-2. PMID: 23670577

Terzi, A., Marinis, T., Zafeiri, A., & Francis, K. (2019). Subject and object pronouns in high-functioning children with ASD of a null-subject language. *Frontiers in Psychology*, *10*, 1301. https://doi.org/10.3389/fpsyg.2019.01301

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV)* [Database record]. *APA PsycTests*. https://doi.org/10.1037/t15174-000

Wiesner, D. (1991). *Tuesday*. Clarion Books.