ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Analysis of Noun Phrase Ambiguity in Narratives Reveals Differences in Referential Establishment But Not Cohesion for Older Autistic Children

Emily Zane[a] and **Ruth B. Grossman**[b]

[a] Department of Communication Sciences and Disorders, James Madison University, Harrisonburg, MI [b] Department of Communication Sciences and Disorders, Emerson College, Boston, MA

ARTICLE INFO

Article History:
Received November 5, 2022
Revision received February 15, 2023
Accepted May 16, 2023

Editor-in-Chief: Stephen M. Camarata
Editor: Mary Alt

https://doi.org/10.1044/2023_JSLHR-22-00630

ABSTRACT

Purpose: Stories told by autistic narrators often contain relatively frequent use of ambiguous references. However, it remains unclear whether this ambiguity is driven by ambiguous character establishment (e.g., "Once upon a time, she/the girl. . .") and/or ambiguous cohesion (e.g., "Two girls lived in a castle. She/The girl. . ."). In this study, we directly compared rates of each type of ambiguity within and between narratives told by autistic and non-autistic children, to determine which type of ambiguity is relatively more common in narratives told by autistic children.
Method: Thirty-three 10- to 17-year-old autistic participants (n = 17) and non-autistic peers (n = 16), who were not statistically different in age, standardized language scores, and IQ scores (p > .8 for all), watched two short animated videos alone and then described the videos' events to two listeners who were openly unfamiliar with the videos. We transcribed video recordings of narratives and coded all referential noun phrases (NPs) as either clear or ambiguous. We further categorized ambiguous NPs as either ineffective introduction or ineffective cohesion.
Results: Autistic children produced significantly higher rates of ambiguous establishment than non-autistic peers, whereas between-group comparisons' rates of ambiguous cohesion were not statistically significant.
Conclusions: Older children on the autism spectrum show differences in the way they introduce characters, selecting NP types that are only appropriate when their listener is already familiar with the referent. In contrast, once they have introduced characters, they show cohesive skills that are comparable to those of non-autistic peers. Findings support theories arguing that autistic children show differences in their application of social pragmatic principles (listener/context-specific pragmatic rules), whereas their use of linguistic pragmatics (context-independent rules) is similar to that of non-autistic peers.

When telling a story, a narrator must effectively manage referential clarity, so their listener can follow each character's actions throughout the story. Referential clarity depends on two main skills. First, the storyteller must establish reference by introducing characters. Second, the teller must maintain referential cohesion by appropriately using pronouns and other noun phrases (NPs) to refer to previously introduced characters.

Establishing reference effectively is largely dependent on the narrator's understanding of a listener's background knowledge. In a story created "from scratch," an appropriate introductory NP is almost always indefinite (e.g., *a man*), unless the story is about the speaker or the listener, in which case deictic[1] pronouns can be used unambiguously

Correspondence to Emily Zane: zaneer@jmu.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

[1] The term "deictic" refers to expressions whose meaning is only interpretable through understanding a speaker's specific spatial, temporal perspective in a given discourse context. For example, in a dialogue, the meaning of the first- and second-person pronouns "I" and "you," respectively, alternates depending on who is speaking.

(e.g., *you* or *I*). However, there are specific circumstances during which definite NPs (e.g., *the man*) can establish reference. Abbott (2004) explains that definite NPs can appropriately introduce a referent when the speaker (a) knows the referent is unique (e.g., "the president of the United States"), (b) knows the listener is already familiar with the referent, (c) assumes the listener can determine reference from context (e.g., "I went to a farm and his job is to feed the chickens"), (d) provides clarifying linguistic information (e.g., "The little boy who lives next door to me is…"; "whoever he is who made the mess in the kitchen…"), and/or (e) uses the term as a deictic and relies on extralinguistic information to clarify the intended referent (e.g., saying "the man on that ladder" when the referent is visible to both hearer and speaker). Similarly, a storyteller can use a third-person pronoun to introduce a character when they know their listener has access to information needed to establish reference. For example, a storyteller could use "she" as a deictic pronoun while relying on extralinguistic cues to disambiguate reference (e.g., the teller points at the intended referent; the teller and listener are both looking at a picture of the intended referent). It is important to again highlight the inextricable link between effectively selecting the form of an introductory NP and having an accurate understanding/representation of a listener's knowledge and perspective; therefore, a narrator's ability to effectively establish characters likely depends on the ability to track and attend to their listener's point of view.

The following provides examples of effective and ineffective introductory NPs:

> Effective referential establishment: Once upon a time, a little girl/the President of the United States/Joe Biden/I…
>
> Ineffective referential establishment: Once upon a time, the little girl/Sarah/she…

Once a storyteller has effectively established reference, they must maintain referential cohesion throughout the remainder of their narrative. This requires appropriately using anaphora: terms that refer back to referents mentioned previously in the discourse. The following examples include anaphoric uses of a pronoun (1) and definite NPs (2) in narrative contexts. In each example, subscripts are used to index the mapping between anaphors and their antecedents.

(1) A little girl named Sally_j lives on a farm. Each morning, she_j gets out of bed very early to feed the chickens.

(2) A little girl named Sally_j and her older sister, Sarah_k, live on a farm. Each morning, Sarah_k/the oldest girl_k gets out of bed very early to feed the chickens.

Referential Ambiguity and Autism

If a story does not successfully establish reference and/or maintain referential cohesion, there will be moments of referential ambiguity. This ambiguity can result from two sources: first, when there is no possible referent because it has not been appropriately introduced (ineffective establishment) and, second, when there are too many possible antecedents (unsuccessful cohesion). We henceforth refer to each type of ambiguity as never-introduced referent and competing referents, respectively. The following provides examples of each.

> Never-introduced–referent ambiguity: Two little boys, Jeremy_j and his older brother Ira_k, live on a farm. Each morning, $\text{she}_?$/the $\text{girl}_?$/$\text{Sarah}_?$ gets up early to feed the chickens.
>
> Competing-referents ambiguity: Two little girls, Sally_j and her older sister Sarah_k live on a farm. Each morning, $\text{she}_{j/k}$/the $\text{girl}_{j/k}$ gets up early to feed the chickens.

Despite both examples resulting in ambiguity, the underlying cause of these mistakes and necessary repair are quite different. The production of never-introduced referential ambiguity may demonstrate a misunderstanding of what background information is necessary before anaphoric NPs can be used and/or an inaccurate understanding of listener background knowledge. To establish reference in such cases, the speaker must provide more context and background information.

In contrast, when producing competing-referents anaphors, the storyteller has provided the listener with necessary background information, but they did not effectively recognize that another antecedent is competing with the target one for the listener's attention. In other words, there are at least two antecedents that are similarly salient to the listener, so that the listener cannot determine which one is referenced by the anaphoric NP. Almor (1999) has explained that antecedent saliency depends on numerous factors, including recency, length, specificity, and grammatical position. In this article, he uses the results from a series of eye-tracking experiments to support the informational load hypothesis, which suggests that there is an inverse relationship between antecedent saliency and anaphor weight. A "heavy" anaphor is one that is more specific (e.g., a definite NP), whereas a "light" anaphor is one that is more general (e.g., a pronoun). Almor's (1999) findings show that listeners process more quickly/easily a light anaphor that refers to a salient antecedent and a heavy anaphor that does not, and any conflict between this expected relationship between antecedent salience and

anaphoric weight results in a processing cost for the listener. In the case of competing-referents ambiguity, the speaker has provided an anaphor that is too light for the given context, not because the target antecedent is not salient but because there is another antecedent that is at least as prominent as the target one. To re-establish referential clarity in such cases, the teller needs to use a heavier anaphoric NP, for example, "the oldest girl" or "Sarah." Therefore, by analyzing the type of referential ambiguity produced by a storyteller, we can determine whether the storyteller has difficulty appropriately tailoring reference to match a listener's background knowledge or effectively tracking the saliency of previous referents.

Referential Ambiguity in Autism

Frequent moments of referential ambiguity, specifically ambiguous pronouns, are commonly observed in narratives told by storytellers on the autism spectrum. Studies find that when both autistic adults (Colle et al., 2008) and autistic children who are between the ages of 8 and 15 years (Banney et al., 2015; Novogrodsky, 2013; Novogrodsky & Edelson, 2016; Suh et al., 2014) generate narratives, they use ambiguous pronouns more often than non-autistic (NA) peers do. In this previous work, autistic and NA participant groups are always similar in age and also show nonsignificant differences in scores on standardized language and/or nonverbal IQ tests.

Surprisingly, even though pronominal ambiguity has been so widely examined in autistic narratives, no previous work has separately analyzed the production of never-introduced–referent and competing-referents ambiguity to determine the underlying reason for referential ambiguity in this population. Colle et al. (2008) focused only on comparisons of never-introduced–referent ambiguity between autistic and NA adults. The majority of the remaining relevant research either combined moments of never-introduced–referent and competing-referents ambiguity to compare the overall frequency of pronominal ambiguity between NA and autistic groups (Banney et al., 2015; Novogrodsky, 2013; Novogrodsky & Edelson, 2016) or did not specify how ambiguity was defined beyond the listener (coder) not being able to disambiguate reference (Baltaxe & D'Angiola, 1996; Norbury & Bishop, 2003; Suh et al., 2014). We are only aware of one previous article that focused on competing-referents ambiguity. They compared ambiguity rates between narratives of 9-year-old autistic and neurotypical children, as well as children with attention-deficit/hyperactivity disorder (ADHD; Kuijper et al., 2015). Interestingly, these authors find that rates of ambiguity are not statistically different between autistic and neurotypical groups, which provides preliminary evidence that autistic storytellers show similar abilities in referential cohesion as neurotypical peers. However, Kuijper et al. (2015) did not directly compare rates of ambiguous introductions (never-introduced–referent ambiguity) versus ambiguous cohesion (competing-referents ambiguity) in their participant population and instead only analyzed the form of introductory NPs (pronouns vs. full NPs). The authors acknowledged that task demands may have been too simple to elicit ambiguity in any participant group, since each story involved only six cartoon panels and contained only two characters, and cartoon panels were visible to participants as they told their stories. Thus, it remains unclear whether competing referents would be more frequent in autistic narratives when they are tasked with more complex stories that require tracking multiple competing characters from memory.

Another limitation of work on referential ambiguity in autistic narratives is the fact that it has almost exclusively focused on the use of ambiguous pronouns, rather than the use of other types of anaphoric NPs, such as definite NPs. Challenges ensuring referential term selection meets the needs of a specific listener and/or tracking referents during discourse should globally affect the use of all referential phrases, not just pronouns. Some evidence for this comes from Tager-Flusberg (1995) and Norbury and Bishop (2003). Although both works only analyzed the ambiguity of pronoun use, each also observed a pattern where a relatively high proportion of autistic participants—but only a minority of NA children—used definite NPs to introduce a novel character (e.g., "Once upon a time, the boy…"). This provides some evidence that autistic children are less likely to use rules related to definiteness and/or their listener's previous knowledge. On the other hand, it is arguable that using definite NPs to introduce characters is not inappropriate when a storyteller dictates a wordless picture book, as was done in both the Norbury and Bishop (2003) and Tager-Flusberg (1995) studies. In such cases, the intended referent is visible to both the listener and the storyteller, so that the referent intended by the definite NP can be resolved through deixis.[2] Still, Norbury and Bishop (2003) comment, more generally, on the "surprising number of ambiguous *nouns* used by [autistic participants, which] was very uncommon in any of the other groups, [and]…requires further attention" (p. 308; italicized emphasis from original text).

Accordingly, there is evidence that autistic individuals show differences in referencing, generally. A relatively recent review summarized findings from 24 articles about reference production by autistic speakers (Malkin et al., 2018). About half the reviewed studies (n = 13) analyzed reference during narrative production, and 11 of these reported that autistic individuals showed relative weaknesses in

[2]This study takes this possibility into account by asking participants to narrate stories about remote actors and events (i.e., events/characters from video shorts they are shown when their listener is not present).

appropriately using referential terms. Results in the two studies not showing this pattern can be explained by task simplicity (Kuijper et al., 2015) and/or inclusion of participants who were young enough (mean age around 7;6 [years;months]) to still be developing mastery of pronoun use (Mäkinen et al., 2014). The patterns of findings from the other 11 narrative studies reviewed in the work of Malkin et al. (2018) are relatively consistent; they show that autistic participants are relatively likely to use pronouns when a full NP would be more effective (Tager-Flusberg, 1995) or vice versa (Arnold et al., 2009; Baltaxe, 1977). Although most of these works did not examine the ambiguity of non-pronominal referential terms, the pattern of difference in referential term selection suggests that it is not the use of pronouns, specifically, that is challenging to individuals on the autism spectrum, but the more overarching skill of selecting the appropriate form of a referential phrase that will effectively identify an intended referent in a particular discourse context.

To summarize, there is a substantial body of evidence suggesting that referential clarity is reduced in stories told by autistic narrators. However, because no previous study has directly compared rates of never-introduced–referent ambiguity to competing-referents ambiguity within and between groups, it is still unclear whether ambiguous referencing is due to an inaccurate representation of the listener's background information or unsuccessful tracking of referents throughout previous discourse. Furthermore, most previous work focuses on pronominal references, meaning that we know very little about how autistic storytellers understand and apply rules for definiteness throughout their narratives.

This Study

This study addresses both gaps by separately coding instances of never-introduced–referent and competing-referents ambiguity for all referential NPs produced in stories told by autistic and NA narrators and then comparing the frequency of each within and between groups. This analysis can help tease apart two proposed explanations for increased rates of referential ambiguity in stories told by autistic narrators: that they (a) do not sufficiently attend to listener's needs (Novogrodsky & Edelson, 2016), perhaps because of differences in how they take into account the listener's point of view (Volden et al., 1997; cf. Beechey, 2022; Milton, 2012), and (b) have "difficulty keeping track of what they…said previously as the story progress[es]" (Norbury & Bishop, 2003, p. 308). The former explanation would predict higher rates of never-introduced ambiguity in stories told by these narrators, whereas the latter would predict increased moments of competing-referents ambiguity.

In this study, we analyze the use of all referential terms (pronouns, nouns, proper names) in two consecutive narratives produced by autistic and NA adolescents and told to two different listeners. We begin this examination by performing a set of foundational analyses to determine whether there are differences between the two groups and the two narratives in the frequency of referential terms, specifically the number of pronouns. We then move on to our primary analyses, which are designed to answer the following research questions: First, are there differences in the relative frequency of never-introduced–referent and competing-referents ambiguity between groups in each of their narratives? Second, are rates of ambiguity differently influenced (between groups) by whether the storyteller has recently described similar events to another person? Our hypotheses to these two questions are as follows. Regarding the first research question, we hypothesize that never-introduced–referent ambiguity will be more frequent in the narratives of autistic individuals as compared to NA peers, based on previous findings. In contrast, based on the findings of the one article that has focused specifically on competing-referents ambiguity in stories told by autistic narrators (Kuijper et al., 2015), we tentatively predict that competing-referents ambiguity will not be more prevalent in autistic narratives than NA narratives. Our hypothesis to the second research question is based on the findings reported in the work of Nadig et al. (2015), who found that autistic individuals were more likely than NA ones to reuse language adapted for a previous listener when speaking to a new one. Accordingly, we predict that autistic participants will show higher rates of ambiguity in the second narrative, as compared to their first narrative, and as compared to NA peers' never-introduced–referent ambiguity rates in either narrative, as autistic participants may be additionally challenged by the task of adapting referential language to meet the needs of a new listener once they have already established reference with someone else.

Method

Participants

We recruited autistic and NA participants, aged 10–17 years, through local schools; advertisements placed in local magazines and newspapers as well as online, including websites for autism advocacy organizations; and word of mouth. During an initial phone interview with participants' caregivers, participants were screened for specific inclusionary criteria: Potential participants in both groups were excluded if they had ever been diagnosed with an intellectual disability, a language impairment, a neurological condition (e.g., epilepsy), or a genetic disorder (e.g., fragile X syndrome), or if they were in special education

classes. We did include potential autistic (but not NA) participants if they had Individualized Education Plans. Potential NA participants were excluded if they had a sibling diagnosed with an autism spectrum condition (n = 2). Potential autistic participants were required to have previously been formally diagnosed with an autism spectrum condition (including pervasive developmental disorder–not otherwise specified and Asperger's).

Participants who met the previous screening criteria were then invited to the lab, where we administered several assessments to ensure typical language and cognitive functioning, confirm autism diagnosis for our autistic participants, and exclude NA participants with social communication differences indicative of autism. To achieve this last aim, caregivers of participants in both groups completed the Social Communication Questionnaire–Lifetime (SCQ-L; Rutter et al., 2003). An SCQ-L score of 15 or higher indicates social communication traits that are associated with autism; therefore, an inclusion criterion for each NA participant was that their SCQ-L score be lower than 15. Participants in both groups also completed the Core Language Subtests of the Clinical Evaluation of Language Fundamentals–Fifth Edition (CELF-5; Wiig et al., 2013) to assess language abilities. We used the Kaufman Brief Intelligence Test–Second Edition (K-BIT-2; Kaufman & Kaufman, 2004) to assess verbal, nonverbal, and combined IQ. For inclusion in the study, participants had to have a minimum standard score of 85 (i.e., more than 1 *SD* below the mean) on both the CELF-5 and K-BIT-2, ensuring that scores were either within or above normal ranges. One potential autistic participant was excluded because their CELF-5 score fell below 85. All autistic participants participated in either Module 3 or 4 (depending on participant age) of the Autism Diagnostic Observation Schedule–Second Edition (ADOS-2; Lord et al., 2012) with a researcher-reliable ADOS administrator, and all autistic participants included in this data set received scores indicative of an autism spectrum condition.

All participants who met these criteria were included and resulted in a final sample of 17 autistic participants (12 boys, five girls) and 16 NA participants (11 boys, five girls). The mean age in years was 13.8 (SD = 2.01) for the autism spectrum disorder (ASD) group and 13.4 (SD = 2.44) for the NA group. See Table 1 for a descriptive summary of demographic information across groups.

Groups did not significantly differ in age, $t(33) = -0.06$, $p = .96$; IQ, $t(33) = -0.03$, $p = .98$; language ability, $t(33) = -0.16$, $p = .88$; or the ratio of boys to girls, $p = 1.00$. As expected, SCQ scores were significantly higher in the ASD group, $t(33) = 8.80$, $p < .001$.

Informed Consent

The institutional review board of Emerson College Boston, Massachusetts, approved this study. We obtained written informed consent from each participant's parent or guardian as well as assent from the participant themselves if they were over the age of 12 years. Participants were compensated for their time with Amazon gift cards; they were paid $15 per hour.

Stimulus Videos

Participants each watched two animated Pixar "shorts" in the same sequence (Lasseter, 1986; Lasseter & Stanton, 1991). Both videos are available on YouTube, and the links to each have been included with their respective listing in the references section.

Both videos involved two animated desk lamps interacting with one another and playing with two balls. In each video, one of the two desk lamps is larger than the other, and each video ends when the smaller lamp makes a foible with one of the balls. While the general characters and events of the two videos are quite similar, there are differences in specific events and in whether the (intact) balls are present on the screen simultaneously. In the first video, one ball deflates before a beach ball is introduced as a replacement. In the second video, the lamps alternate between engaging with a beach ball and a bowling ball, meaning that both the beach and bowling balls are present on the screen at once.

Table 1. Summary of demographic information for each group and between-group comparisons.

Information	Autistic (n = 17)	NA (n = 16)	Test statistic	p
Age	13.79 ± 2.24	13.83 ± 1.98	t = −0.06	.96
Sex (F:M)	5:12	5:11	N/A (Fisher's exact test)	1.00
IQ	109.82 ± 18.51	110.00 ± 18.41	t = −0.03	.98
Language	106.71 ± 15.72	107.69 ± 19.45	t = −0.16	.88
SCQ	19.47 ± 7.24	3.00 ± 2.61	t = 8.80	< .001

Note. NA = non-autistic; F = female; M = male; N/A = not applicable; SCQ = Social Communication Questionnaire.

In both videos, the audio includes the sounds of the lamps and balls moving around, along with piano music playing in the background. The second video originally contained a bit of narration (a man's voice is twice played saying the word "heavy," and a woman's voice is twice played saying the word "light"). This narration was spliced out of the audio track and replaced with audio that came immediately before/after the word was played, so that the final audio track contained uninterrupted piano music and the sounds of the lamps/balls moving, without any narrated words. Additionally, the title screens and final credits were removed from both videos.

Procedure

This experiment occurred at the beginning of a 2-hr research visit to the lab. At the start of this experiment, a researcher (Researcher A) brought the child into the testing space and asked the child to sit at a chair in front of a computer monitor. An HD camera on a tripod was positioned facing the chair, and the camera was turned on before the participant entered the room, so that the child was recorded from the moment they sat in the chair. While they were video-recorded, participants watched two short videos. After each video, the participant was asked to describe the events of each video to a different listener, Researcher A for Video 1 and Researcher B for Video 2, which resulted in two narratives per child: Narrative 1 and Narrative 2. Both Researchers A and B were NA adults. Detailed descriptions of the procedures used to elicit each narrative are described in the First Narrative and Second Narrative subsections of the Procedure section.

Our narrative elicitation task includes several modifications to experimental procedures that have traditionally been used to measure autistic individuals' use of referential terms. First, we asked children to narrate events strictly from memory, rather than allowing them to rely on visual cues as they produced their story. This manipulation was important as it prevented deictic uses of definite NPs or pronouns (e.g., saying "he" or "that guy" to refer to a character visible to both storyteller and listener). It also increased the cognitive demands of our task, as participants had to rely on their memory to accurately present story events and characters. We hoped this would prevent ceiling effects in both groups and could amplify group differences that were not evident when task requirements were too simple (e.g., Kuijper et al., 2015). A manipulation that additionally increased task demands for the second narrative was that participants needed to describe/differentiate the storylines and characters of two animated shorts, when both shorts involve similar characters and events. Therefore, when telling the second narrative to the second listener, participants had to ensure that they did not confuse events between the two shorts. Finally, both animated shorts included four characters that appear at different points in the film, who could each be referenced using similar NPs, including the pronoun "it." This final manipulation demanded that participants had to introduce multiple referents at varied points in the story, rather than only one character at the very start of the story, and participants also must track multiple (potentially competing) antecedents throughout the story. By doing so, we provided many opportunities for the two types of ambiguity we would code and analyze.

We further increased task complexity by asking children to adjust referential term selection to meet the needs of two different naïve interlocutors, requiring participants to introduce characters to one naïve interlocutor and then to start again (reintroduce the same characters) when they describe film events to a new naïve listener. We call our paradigm a "narrative relay," as the two interlocutors interact with the children as a "tag team," each listening to one of the child's narratives, requiring the child to tell two narratives in a sequence: Participants told the first narrative to the first interlocutor immediately after watching one video and the second narrative to the second interlocutor immediately after watching another. Only one previous study used a similar manipulation, where autistic and NA adults were tasked with establishing referential terms for objects with one partner, who was then eventually replaced with someone new (Nadig et al., 2015). These authors found that autistic adults were able to adjust their referential terms to the first partner, but they were marginally less likely than NA individuals to appropriately adapt their language when that partner was replaced with someone else.

First Narrative

Researcher A briefly left the room, and when she returned, she informed the participant that there was a problem with equipment needed for the first experiment. She told the child that another volunteer in the lab had saved a video onto the computer for "just this kind of situation." She said that she had never seen the video before, but she trusted that the volunteer picked something good. She turned on the video and promptly left the testing room, shutting the door behind her. As soon as she left the room, she started a timer that was preset for the video's duration. Meanwhile, in the testing room, the child watched the video by themselves. The video started with 30 s of silence alongside a black screen, allowing time for the researcher to leave the room and shut the door before any content played.

After the timer indicated that the video had ended, Researcher A returned to the room. As she entered, she said, "I want to hear about that video you watched. Like I said, I've never seen it before." Thus, every participant

had two opportunities for learning that Researcher A was not familiar with the video: once before the participant watched the video and once after. Researcher A sat in a chair opposite the child and asked, "What was that video about?" If the child proceeded in telling a complete narrative (all the video's events from start to end), the researcher only provided back channeling (e.g., nodding, saying "Uh huh") and other types of generic comments (e.g., "Oh really?") as the child talked.

If a child did not initiate a narrative on their own, Researcher A would ask a series of standardized questions to prompt the child to tell a complete narrative capturing all the video's events: "What happened at the beginning of the video?", "And then what happened?", and "What happened at the end?" The researcher repeated, "And then what happened?" as many times as needed to ensure the child described all the events in the story. If the child skipped an event in their story, Researcher A would ask a question to explicitly target that event, for example, "What happened after the ball deflated?" Once the child finished telling the entire video's story, Researcher A asked the child what they thought of the video—whether they liked it and whether the research team should show it to other participants if there was another unexpected delay between tasks. Most children said that they enjoyed the video.

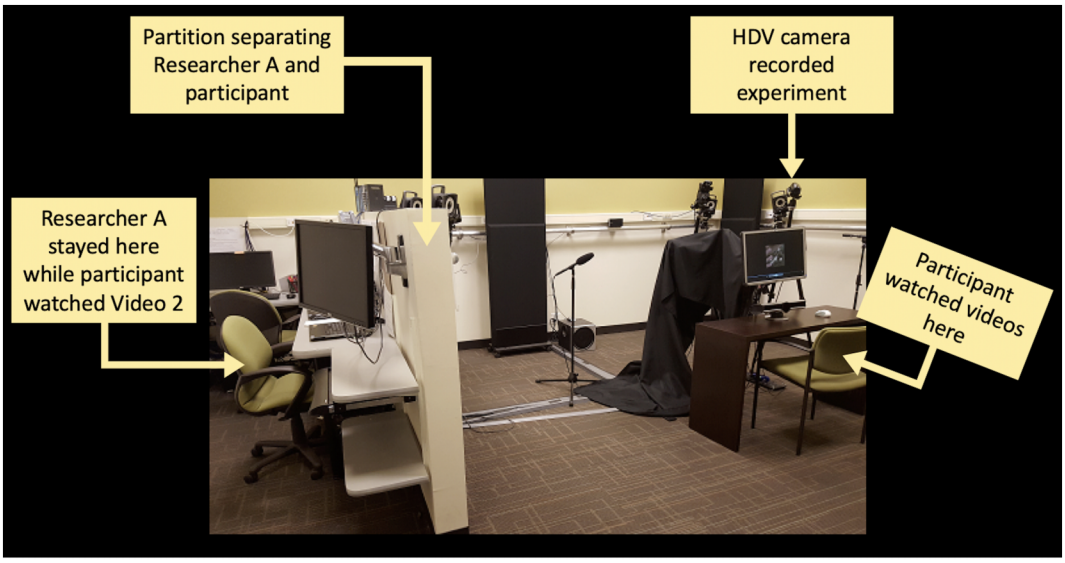
Second Narrative

After Researcher A ensured the child completed telling the entire first video's narrative, she told the child that she assumed the equipment was still not ready, as Researcher B had not yet entered the testing room. She informed the participant that they both (Researcher A and the participant) must continue to wait. Researcher A then suddenly "remembered" that the same volunteer who saved the lamp/ball video on the computer in front of the child had also saved a second video on that computer, so she told the participant that they should watch that video while they wait. Researcher A started the second video and then went behind a partition (see Figure 1) and sat down so that the child could not see her. The child watched the video while Researcher A remained behind the partition. Researcher A stayed silent during the duration of the video, even if the child attempted to initiate an interaction with her.

As soon as the video ended, Researcher A stood up, walked to the door of the testing room, and opened the door. Meanwhile, Researcher B was waiting outside the door. Loudly, so that the participant could hear, Researcher B told Researcher A that there was still a problem with the equipment, and then Researcher B asked Researcher A to examine the equipment. Researcher B then entered the testing room and closed the door behind her, leaving Researcher A outside the testing room behind a closed door. This marked the "baton-passing" moment in the "relay" paradigm, where Researcher A passed on the responsibility of the task to Researcher B.

Researcher B then sat down on the chair opposite the child and apologized for the delay. She asked the child what the child and Researcher A had been doing while waiting. Children all responded with some variation of, "Watching videos." Researcher B initially feigned confusion and then said something such as, "Oh! I know what

Figure 1. Position of Researcher A and participant as the participant watched the second animated short.



videos you mean. A volunteer saved them on that computer. I've never seen them before. Did you just finish watching one? What was it about?" The procedures in Interview 2 were like those in Interview 1, except that Researcher B was responsible for ensuring that the child focused only on telling the events of the story from the second video. Researcher B used the same series of prompts Researcher A used to elicit the full story of the second video.

Once the story was complete, Researcher B stood up, walked to the door of the testing room, opened it, and called out to Researcher A, asking if the equipment was finally working. Researcher A said yes, entered the room, and the researchers began another study. The entire process of watching both videos and conducting the two interviews took about 10 min (M = 9:52, SD = 2:18), beginning from the moment the child started watching the first video to the conclusion of the second interview.

Participant Deception

As is clear in the preceding sections, this experimental paradigm included mild deception, in which participants were made to believe that (a) Researchers A and B were completely naïve to the content of the animated shorts and (b) they were watching videos before experimental tasks began, rather than as part of an experimental task. While the inclusion of participant deception is never ideal, we decided it was necessary for addressing our aims.

Regarding the first deception: To analyze never-introduced ambiguity, specifically, it was important that participants were told that their listeners had no previous knowledge of story characters and events. If it was not made clear that listeners had no knowledge of video events, then ineffective reference establishment could be attributed to a participant assuming the researchers already knew the characters being referred to. Furthermore, by telling participants that their listeners had no previous knowledge of the videos, we hoped to (implicitly) encourage them to be clear, and to tell stories in detail, because they could not rely on their listener to "fill in the gaps" based on background knowledge.

Regarding the second deception, we wanted participants to produce narratives about past events like they would in the "real world," without the added pressure of knowing that their stories would be analyzed for clarity and cohesion. In this way, we could more comfortably generalize findings about referential clarity to storytelling events outside of an experimental/laboratory context. Prior to completing this task, all participants provided informed consent, which included telling them that they would be video- and audio-recorded the entire time they were in the experimental room. The video camera recording this experiment was powered on (with a light indicating this was the case) and was directly in front of participants while they watched videos and told stories (see Figure 1).

We debriefed experimental procedures, including this one, with participants and their caregivers at the end of the 2- to 3-hr research visit.

Transcription

Videos of each child's production of both narratives were transcribed in ELAN (Brugman & Russel, 2004; ELAN, 2022) by at least two research assistants (RAs), who were unaware of the diagnostic status of the child in the video. One RA was responsible for creating a first-pass transcription of all the language in the video (including Researcher A and B's speech). All speech was transcribed in capital letters, and utterance boundaries were demarcated by the creation of separate annotations per utterance. Periods and commas were not used, but question marks were used to indicate question intonation, and ellipsis was used to indicate a short pause that occurred within utterance boundaries.

Once the first RA finished transcribing a participant's narratives, another RA was tasked with completing a second pass. Second-pass transcribers watched each narrative with the ELAN transcript open and read through the transcription while listening to the narrative and looking at participant/researcher gestures. If the second-pass transcriber did not dispute any of the transcription, the file's transcription was determined final and ready for coding. This happened rarely (n = 2). For most files (n = 31), the second-pass RA disagreed with some portion of the initial transcription, and they would indicate this in the transcription file by using bracketing to indicate their suggested change to the segment in question (e.g., THEN THE [THIS] OTHER LAMP ENTERED). When there was any such disagreement between the first- and second-pass RAs, the file was reviewed by a third RA (third-pass transcription), who was responsible mainly for examining segments of the transcription where there was disagreement and for selecting the transcription that they believed was most accurate. If the third-pass transcriber was able to settle all disputes, the file was deemed final (n = 16). In cases where the third-pass transcriber could not choose between either option—either because both seemed equally possible or because neither seemed correct—the file moved onto consensus transcription (n = 15). Files were also moved to consensus transcribing if the third-pass transcriber noticed an error in the initial transcription that had not been caught by the second-pass transcriber. Consensus transcribing required that all three transcribing RAs met, watched the relevant portion(s) of the video together, and made a final determination as a group.

Analysis

Coding

Final transcriptions of each narrative were then coded for the clarity of referential NPs. A team of six RAs—unaware of participants' diagnostic status—were trained to code narratives for referential clarity, where three RAs were each responsible for independently coding every narrative. Coding, like transcription, was completed in ELAN. Coders first read through each child's transcription and identified all the NPs that referred to either of the two lamps or either of the two balls in each video. There were many types of referential NPs that were eventually identified: pronouns (e.g., "it," "they," "he"), determiner–noun combinations (e.g., "a smaller lamp," "that first lamp," "a ball"), proper names (e.g., "Luxo"), and larger phrases that included, for example, prepositional or relative-clause complements (e.g., "one of them," "the ball that he had at the beginning"). Finally, coders focused on the words that referred directly to the lamps or balls and coded these for clarity. For example, in the phrase "the ball that he had at the beginning," only the smaller NPs referring directly to the ball ("the ball") and the lamp ("he") would be coded. The rest of the phrase ("that…at the beginning") was used to determine whether the referents for these two phrases were clearly identifiable.

Once referential NPs were identified, coders labeled their lexical category (e.g., pronoun, determiner, noun) and then determined their clarity, as follows. First, NPs were coded as clear or ambiguous. Clear NPs were those for which coders could establish the intended referent. This occurred either because the phrase clearly referred to a character/object that had already been introduced or because the NP introduced a character/object appropriately (e.g., "at the beginning of the video, there was a lamp by itself"). Ambiguous NPs were those where coders could not establish reference. Once NPs were coded as ambiguous, coders selected the reason why the phrase was ambiguous from a drop-down menu (never-introduced referent or competing referents). Never-introduced–referent ambiguity was selected when reference could not be established for a given NP because the character/object being referred to has not yet been introduced (e.g., using a definite article—"the lamp" or "the ball with red stripes"—in referential NPs at the very beginning of the narrative).[3] Competing-referents ambiguity was coded when reference could not be established for a given NP because it was unclear to which object/character mentioned recently in the discourse the NP referred (e.g., "the two lamps hit the ball back and forth, but then he started jumping on it").

In NPs with determiners, both the noun itself and the determiner were coded in the same way, so if the NP "the lamp" was determined to be ambiguous due to its referent never having been introduced, both "the" and "lamp" were coded for never-introduced ambiguity. This means that there is some overlap in measuring determiners and nouns that are ambiguous. However, we elected to include them separately as there were 59 nouns (of 1,097 total) that were produced without determiners (including plural nouns and proper names). Including determiners and nouns separately allowed us to capture these 59 nouns.

Reliability

Once all files were coded, codes were compared for agreement using a script created in RStudio (RStudio Team, 2022). This script compared coders' determinations of the following: which words should be coded in the first place (which words made up referential phrases), coders' determinations of lexical category (pronoun, noun, determiner, etc.), and coders' determination of clarity (clear, never-introduced ambiguous, competing ambiguous). When the script determined that at least two of three coders agreed upon all three designations for a single word, that word's coding was deemed complete, and final codes reflected those selected by the majority. As an example: Imagine all three coders elected to code the word "he"; one coder said it was a noun, whereas the other two coded it as a pronoun, and one coder coded it as ambiguous competing, whereas the other two coded it as never-introduced ambiguous. In this case, "he" would be coded as a referential pronoun that was ambiguous because its referent had not been introduced.

In some cases, a majority decision was not possible for a word or an entire phrase. For example, one coder may have determined the word "that" was an ambiguous (competing-referents) pronoun, another as a clear determiner, and a third as an ambiguous (never-introduced–referent) determiner. In those cases, that segment of the transcription was coded by a fourth coder, who had never coded that transcription, and their codes were compared to the initial three codes. Out of the 4,290 words coded as referential, such disagreements were rare; there were 29 of them in total (n = 21 disagreements for clarity/type of ambiguity [0.49% of words], n = 8 disagreements for part of speech [0.19%]). After the fourth coder completed these 29 codes, the fourth coder's codes were again compared to Coder 1–3's codes. When two of four coders agreed, this agreed-upon code was included in the final set of data, so

[3]It is important to note that our definition of never-introduced–referent ambiguity is stricter than in many previous studies, as we only coded a phrase ambiguous when its referent had never been introduced, whereas other studies (e.g., Colle et al., 2008) also coded NPs as ambiguous when an intended antecedent was "too distant" from an anaphoric referent.

that after this step, the codes for all 4,290 referential words reflected codes that were assigned to the word by at least two different people.

Statistical Comparisons

Before comparing rates of ambiguity between groups, we summed the number of referential terms per participant per narrative and used a 2 × 2 (Group × Narrative) analysis of variance (ANOVA) to compare the total amount of referential language used in each narrative between groups and to compare the number of pronouns used, as previous research has suggested that autistic storytellers use fewer references, and particularly pronominal references, than NA storytellers (Baltaxe et al., 1995). To account for such differences, we elected to calculate proportions of ambiguous terms out of the number of total referential words (per participant) and use those proportions in statistical comparisons of ambiguity-type analyses, rather than comparing raw frequencies of ambiguous term use. We used a 2 × 2 × 2 repeated-measures ANOVA to address Research Questions 1 (Is there a difference in rates of different types of ambiguity between groups?) and 2 (Are rates of ambiguity affected by narrative sequence?). In this ANOVA, fixed predictors were Group (Autistic vs. NA), Type (Never Introduced vs. Competing), and Narrative Sequence (First vs. Second). The first two predictors were included to answer Research Question 1; the last predictor was added to answer Research Question 2.

Finally, as a post hoc measure, we used a 2 × 2 × 3 repeated-measures ANOVA to investigate whether certain lexical categories were more likely to be ambiguous, generally; whether this differed by subtype of ambiguity; and whether this differed between groups. The predictor variables for this ANOVA were Group (Autistic vs. NA), Ambiguity Subtype (Never-Introduced–Referent vs. Competing-Referents), and Lexical Category (Nouns vs. Determiners vs. Pronouns). We did not include a predictor variable of Narrative Sequence (First vs. Second) in this analysis, as we had no predictions as to whether a certain lexical category would be more frequently ambiguous in one narrative versus the other.

Results

Referential Term Use

An initial 2 × 2 ANOVA was used to determine whether the total number of terms used to refer to the lamps and balls differed between groups and/or between the two different narratives (First vs. Second). There was a significant effect of narrative sequence, $F(1, 31) = 5.13$, $p = .03$, $\eta^2G = .03$, reflecting the fact that both groups used more referential terms in the second narrative compared to the first. There was no significant effect of diagnostic group, $F(1, 31) = 0.05$, $p = .82$, $\eta^2G = .00$, nor was the interaction between Group and Narrative Sequence significant, $F(1, 31) = 1.28$, $p = .27$, $\eta^2G = .01$. Table 2 includes averages and standard deviations for the number of referential terms used by each group, overall, and within each narrative, along with group comparisons.

To determine whether there was a difference in the number of pronouns used between autistic and NA groups (e.g., Arnold et al., 2009), we used a repeated-measures 2 × 2 (Group × Narrative) ANOVA with the number of pronouns as the dependent variable. There were no significant main effects or interactions (Group: $F(1, 31) = 0.05$, $p = .83$, $\eta^2G = .00$; Narrative: $F(1, 31) = 0.51$, $p = .48$, $\eta^2G = .00$; Group × Narrative: $F(1, 31) = 1.76$, $p = .20$, $\eta^2G = .01$). See Table 2 for averages and standard deviations.

Ambiguous Term Use

We used a 2 × 2 × 2 ANOVA with Group (Autistic vs. NA), Narrative (First vs. Second), and Ambiguity Subtype (Competing Referents vs. Never-Introduced Referent)

Table 2. Average number of referential terms and pronouns used by each group and within each narrative.

Variable	Autistic	Non-autistic	Test effect	F	*p*	η^2G
Referential terms						
Total	65.74 (32.17)	63.75 (22.27)	Group	0.05	.82	.01
1st narrative	63.30 (26.60)	56.56 (11.75)	Narrative × Group	1.28	.27	.00
2nd narrative	68.18 (37.61)	70.94 (27.88)				
Pronouns						
Total	38.88 (20.02)	40.13 (14.47)	Group	0.05	.83	.01
1st narrative	19.88 (9.00)	18.50 (5.33)	Narrative × Group	1.76	.20	.00
2nd narrative	19.00 (12.82)	21.69 (10.42)				

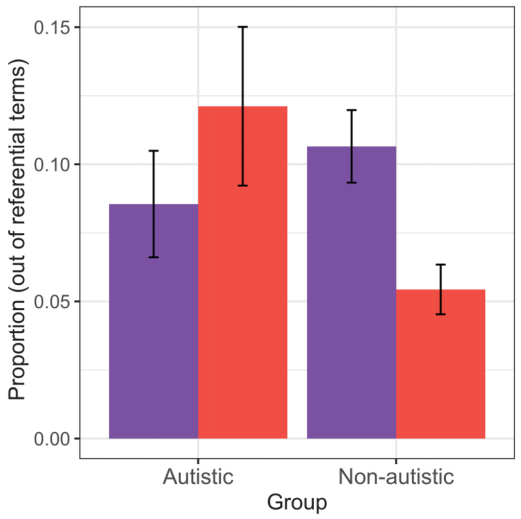
Note. Test statistics include all group (autistic vs. non-autistic) effects and interactions from 2 × 2 (Group × Narrative) repeated-measures analyses of variance.

as predictors of the proportion of ambiguous terms out of the number of referential terms. For this analysis, we elected to use proportions of ambiguous terms—rather than the raw number of ambiguous terms—because, as reported in the previous section, there was a significant difference in the total number of referential terms used between the two narratives in both groups, where both groups of participants used significantly more referential terms in the second narrative as compared to the first (p = .03). Since every time a participant uses a referential term, it has the potential to be ambiguous, an increased use of referential terms could yield a higher number of ambiguous references, even when the relative rate of ambiguity remains the same. By using proportions in our analyses, we ensured that any differences in the occurrence of ambiguity between the two narratives was capturing differences in the relative frequency of ambiguity, rather than differences in narrative length or number of referents produced.

The ANOVA yielded no significant main effects (Group: $F(1, 31) = 1.28$, $p = .27$, $\eta^2G = .02$; Narrative: $F(1, 31) = 0.15$, $p = .70$, $\eta^2G = .00$; Ambiguity Subtype: $F(1, 31) = 0.15$, $p = .70$, $\eta^2G = .00$). However, the pertinent interaction for Research Question 1—the interaction between Diagnostic Group and Ambiguity Subtype—was significant, with a small effect size, $F(1, 31) = 6.17$, $p = .02$, $\eta^2G = .05$, reflecting that a larger proportion of referential terms produced by the autistic group were coded as yielding never-introduced ambiguity (compared to the NA group and compared to competing ambiguity; see Figure 2).

All interactions pertinent to Research Question 2 (whether narrative sequence differentially impacted ambiguity rates, overall, or ambiguity subtypes for each group) were not statistically significant (Group × Narrative: $F(1, 31) = 0.93$, $p = .34$, $\eta^2G = .01$; Narrative × Ambiguity Subtype: $F(1, 31) = 1.86$, $p = .18$, $\eta^2G = .01$; Group × Narrative × Ambiguity Subtype: $F(1, 31) = 0.40$, $p = .53$, $\eta^2G = .00$). Table 3 presents the number and proportion of ambiguous terms (out of referential terms) for each group within each narrative, as well as test statistics and p values for all comparisons involving diagnostic group.

Figure 2. Average proportions of referential terms for each group that are ambiguous by subtype. Purple: competing-referents ambiguity; red: never-introduced–referent ambiguity. Error bars indicate standard error.



Because the interaction between Diagnostic Group and Ambiguity Subtype was significant, we followed the ANOVA with a post hoc Tukey's honestly significant difference test, which yielded a significant difference between proportions of never-introduced–referent ambiguous terms out of total referential terms between groups (p = .05), reflecting the fact that a higher proportion of the referential terms used by autistic children referred to never-introduced characters, compared to NA peers (see Figure 2). There was not a significant difference in proportions of referential terms that were ambiguous due to competing referents (p = .61), nor were there significant differences in proportions of ambiguity type out of total referential terms within the autistic group (p = .33). Within-group comparisons showed a marginal difference for the NA group (p = .08), reflecting marginally higher proportions of their referential terms that were coded as having competing referents (vs. having never-introduced–referents) ambiguity in their narratives.

Lexical Category

We used a 2 × 2 × 3 ANOVA to analyze how Group (Autistic vs. NA), Ambiguity Code (Never-Introduced Referent vs. Competing Referents), and Lexical Category (Pronoun vs. Noun vs. Determiner) impacted proportions. The main effect of lexical category was significant, $F(2, 62) = 17.54$, $p < .001$, $\eta^2G = .06$, reflecting that proportions of pronouns that were coded as ambiguous (out of the total number of pronouns used) were higher across groups (Autistic: M = 13.86%, SD = 14.75%; NA: M = 11.75%, SD = 11.45%), compared to determiners (Autistic: M = 11.11%, SD = 14.22%; NA: M = 7.05%, SD = 5.94%) and nouns (Autistic: M = 8.14%, SD = 10.77%; NA: M = 5.79%, SD = 5.82%). The interaction between Lexical Category and Ambiguity Type was also significant, $F(2, 62) = 48.91$, $p < .001$, $\eta^2G = .25$, reflecting an across-group pattern where pronouns were more often coded for ambiguity due to competing antecedents rather than because their referent had been introduced, with the opposite being true for nouns and determiners (see Figure 3).

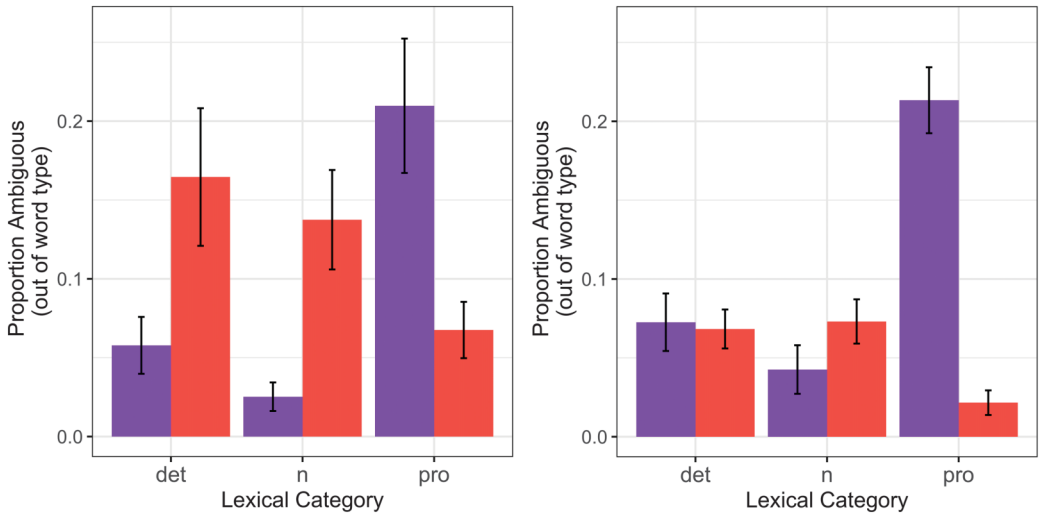
No interactions involving Group and Lexical Category were significant, nor was the main effect of Group or Ambiguity Type, but the interaction between Group and Ambiguity Type was significant, $F(1, 31) = 5.09$, $p = .03$,

Table 3. Average number and percentages (*SD*s in parentheses) of ambiguous references for each group within each narrative.

Variable	Autistic	Non-autistic	Test effect	*F*(1, 31)	*p*	η^2G
Ambiguous terms						
Total	13.56/65.74 (10.70/32.17) 20.67% (17.21%)	9.81/63.75 (5.88/22.27) 16.09% (9.40%)	Group	1.28	.27	.02
Never-introduced referent	7.15/65.74 (6.19/32.17) 12.12% (14.22%)	3.25/63.75 (3.67/22.27) 5.45% (5.75%)	Group × Ambiguity Subtype	6.17	.02*	.05
Competing referents	6.41/65.74 (6.57/32.17) 8.55% (7.88%)	6.56/63.75 (4.01/22.27) 10.65% (7.22%)				
Total ambiguous terms within narrative						
1st narrative	12.71/63.29 (11.20/26.60) 19.90% (15.47%)	10.00/56.56 (5.63/11.75) 18.03% (10.36%)	Group × Narrative	0.93	.34	.01
2nd narrative	14.41/68.18 (10.45/37.61) 21.43% (10.36%)	9.63/70.84 (6.30/27.88) 14.15% (8.20%)				
Ambiguous subtypes within narratives						
Never-introduced referent						
1st narrative	6.47/63.29 (5.52/26.60) 11.24% (10.06%)	2.88/56.56 (2.99/11.75) 5.08% (5.37%)	Group × Ambiguity Subtype × Narrative	0.40	.53	.00
2nd narrative	7.82/68.18 (6.90/37.61) 13.00% (17.72%)	3.63/70.84 (4.32/27.88) 5.79% (6.26%)				
Competing referents						
1st narrative	6.24/63.29 (7.50/26.60) 8.66% (8.21%)	7.13/56.56 (4.49/11.75) 12.95% (8.35%)				
2nd narrative	6.59/68.18 (6.90/37.61) 8.44% (7.79%)	6.00/70.84 (3.52/27.88) 8.35% (5.19%)				

Note. Proportions represent numbers of each subtype of ambiguous references (numerator) out of the total number of referential terms used for each group (denominator). Denominators for test effects including Narrative as a variable represent number of referential terms used by each group within each narrative. Test statistics represent values from 2 × 2 × 2 (Group × Ambiguity Type × Narrative) repeated-measures analyses of variance.

Figure 3. Average proportions of ambiguity subtypes for determiners (det), nouns (n), and pronouns (pro). Purple: competing-referents ambiguity; red: never-introduced–referent ambiguity. Error bars indicate standard error. Left plot: autistic; Right: non-autistic.



$\eta^2G = .05$, as it was in the 2 × 2 × 2 ANOVA described in the previous section.

Discussion

Findings yield three major conclusions. First, ambiguous referential terms are not relatively more frequent in the narratives of autistic children compared to NA peers, at least not for this set of participants, that is, 17 autistic and 16 NA adolescents (10–17 years old), with standardized language and IQ scores in typical ranges. Our second major finding is that, although this group of autistic children was not more ambiguous overall, they did use proportionally significantly more terms that are ambiguous due to their referent never having been introduced, compared to NA peers. There were no differences in the number of competing-reference ambiguity between groups, nor were there significant differences in the type of ambiguity used within groups. Thus, never-introduced ambiguity was the only type of ambiguity that was more prevalent in the narratives told by this set of autistic children. The third major finding is that while both groups of children are more ambiguous when they use pronouns than when they are using nouns or determiners, it is not pronouns that drive the relatively high proportions of never-introduced referential terms in the autism group. We discuss all three of these findings in more detail in the following sections.

Overall Ambiguity

We do not find higher rates of referential ambiguity in the autism group, overall. In fact, children in both groups used referential terms that were ambiguous more than 15% of the time, suggesting that ambiguous referencing is not uncommon, nor is it unique to narratives told by autistic individuals. We highlight this finding as revealing strengths in referential clarity for our autistic participants, a finding that lies in stark contrast to many previous studies, which reported significantly higher rates of pronominal ambiguity in the narratives of autistic participants compared to NA participants (Banney et al., 2015; Colle et al., 2008; Norbury & Bishop, 2003; Novogrodsky, 2013; Novogrodsky & Edelson, 2016; Suh et al., 2014). We offer two main explanations for differences between our findings and those of previous work.

First, our narrative prompts may have encouraged more ambiguity than those used in previous studies. In the current paradigm, children were tasked with initiating and maintaining referential clarity while telling narratives that involved more than one character (i.e., two lamps in both videos) interacting with one another and with two additional objects (i.e., two balls). Crucially, in the current paradigm, characters and objects could be referenced with identical labels. Not only could the phrases "the lamp" and "the ball" each refer to two different referents, but also, children could use the pronoun "it" to refer to four different entities in each video. Thus, participants had to be very careful in their choice of referential phrases to ensure that when they used the pronoun "it" or even a full NP like "the lamp," it was clear from the immediate context which lamp was being referenced, and by using clarifying modifiers when necessary (e.g., "the smaller lamp").

In contrast, many previous paradigms that report increased rates of ambiguity in narratives told by autistic children have used *Frog, Where Are You?* (Meyer, 1969) as their story prompt (Colle et al., 2008; Norbury & Bishop, 2003; Novogrodsky, 2013; Novogrodsky & Edelson, 2016), and others (Banney et al., 2015; Suh et al., 2014) have used *Tuesday* (Weisner, 1991), a story that is employed during administration of the ADOS-2. Both stories involve multiple characters, but the task of clearly referencing them is more straightforward than it is in our current paradigm. In the case of *Frog, Where Are You?*, each character is a unique animal/person (e.g., one boy, one dog, one frog), so that full NPs (e.g., "the boy") can only refer to one referent. Similarly, the plot of *Tuesday* involves a group of frogs interacting with various unique characters on each page (e.g., a dog, a woman), which can also be referenced unambiguously using full NPs. In addition, although *Tuesday* involves multiple frogs, there is no main character frog to whom a storyteller would need to refer contrastively. Instead, the frogs act as a group, reducing the need to track and maintain clear reference for any individual character.

The second related reason our study may not have yielded group differences in overall rates of ambiguity is that our narrative prompts involved multiple characters to whom similar language could refer. Therefore, our prompts encouraged a relatively high rate of competing-referents ambiguity, which we found was equally prevalent in the narratives of both groups. Because previous studies used narrative prompts that did not allow for much competing-referents ambiguity, and because they collapsed their data across both competing-referents and never-introduced–referent ambiguity when comparing groups (Banney et al., 2015; Novogrodsky, 2013; Novogrodsky & Edelson, 2016) or did not specify their definition of ambiguity at all (Baltaxe & D'Angiola, 1996; Norbury & Bishop, 2003), it is possible that group differences identified in those studies were attributable specifically to higher rates of never-introduced–referent ambiguity in the autistic group, which is the same pattern found in our study. The findings in the work of Colle et al. (2008) provide some confirmation for this explanation. These authors only coded and compared never-introduced–referent ambiguity

between groups, and they identified a significant group difference. This indicates that never-introduced–referent ambiguity specifically drives overall group differences identified in previous work.

Mäkinen et al. (2014) also reported similar rates of ambiguity between groups. They explain this pattern as stemming from the fact that the Finnish pronoun *se* can refer to all manner of referents, including inanimate objects. They further suggest that the NA participants in their study (who were 7 years old on average) were still developing unambiguous use of this pronoun. The fact that we replicate their findings with older autistic and NA children suggests that all people may be encouraged to produce high rates of competing-referents ambiguity when they tell a story involving several characters that their language system can reference in the same way.

Referential Ambiguity Subtypes

Our study found that autistic children produce significantly more terms that were ambiguous due to having a never-introduced referent than NA children. There were no between-group differences in the frequency of competing-reference ambiguity, indicating that autistic participants in this study were just as competent as their NA counterparts at using anaphoric references to maintain referential cohesion throughout their narratives. This ability, although captured by the findings of some previous work (e.g., Kuijper et al., 2015; discussed in more detail below), has not previously been highlighted as an area of proficiency for autistic storytellers.

In contrast, autistic participants showed significantly higher proportions of never-introduced–referent ambiguity. Never-introduced–referent ambiguity is listener dependent, in that referencing "the lamp" is only ambiguous in cases where a particular listener is unfamiliar with the target referent. In contrast, competing-referents ambiguity is not listener specific. That is, using the pronoun "he" in a story involving several, salient reoccurring male characters is ambiguous regardless of who the listener is or what background knowledge they have. Thus, never-introduced ambiguity entails differences in recognizing the listener's knowledge or, in this case, the lack thereof. This could explain why autistic children produce proportionally more never-introduced ambiguity than their NA counterparts: They specifically struggle to clarify reference to someone who does not share their background knowledge. This accords with the results from a recent study indicating that autistic children show differences in their receptive interpretation of "social" pragmatics, whereas their grasp of "linguistic" pragmatic principles is similar to that of NA peers (Andrés-Roqueta & Katsos, 2020). According to these authors, "social pragmatics" refers to pragmatic rules that rely on forming, maintaining, and updating accurate representation of someone else's knowledge/ideas/feelings, whereas "linguistic pragmatics" refers to fixed pragmatic rules (i.e., Gricean maxims) that do not require "on-the-spot reasoning about speaker's intent" (p. 1496). This study replicates these findings with expressive data, showing that older autistic children do not show difficulties with the fixed linguistic rules dictated by the informational load hypothesis (Almor, 1999). Thus, they are as good at providing appropriately specific NPs when there are multiple competing referents as their NA peers. However, when they must adapt their message to meet the specific social demands of a particular storytelling context and naïve listener, they are prone to producing more ambiguous terms than NA peers.

These findings may also provide support for claims that ambiguity in narratives told by autistic storytellers is caused by challenges in meeting listener needs (e.g., Novogrodsky & Edelson, 2016), and they may additionally cast doubt on a secondary explanation for increased ambiguity—challenges with executive functioning. Some previous work has proposed that as autistic individuals tell stories, previous mentions of a referent may be less stable in their working memory, which can lead to two patterns of usage: (a) using a pronoun when the referent is too distant to be retrievable (Colle et al., 2008; Novogrodsky, 2013; Novogrodsky & Edelson, 2016; Tager-Flusberg, 1995) and (b) using full NPs when pronouns are possible (Arnold et al., 2009; Baltaxe, 1977). In support of this hypothesized relationship between referencing and executive functioning, Arnold et al. (2009) found a significant relationship between narrative disfluency, which they interpreted as an indicator of cognitive load, and referential form selection for all autistic and NA children, suggesting that problems with appropriate reference selection during narrative production may indeed stem from the cognitive challenge of tracking referents while telling a story.

Although the findings from the work of Arnold et al. (2009) do suggest that executive functioning may play a role in what anaphoric form a speaker selects to maintain referential cohesion, our data suggest that differences in executive functioning may not underlie increased rates of referential ambiguity for autistic storytellers. Beginning with the type of ambiguity that was most prevalent in their stories, never-introduced–referent ambiguity, our coding scheme only included NPs for which a referent truly had not been mentioned, versus NPs where a previous mention was "too distant" (Colle et al., 2008) for the listener to retrieve. Consider the following excerpt from the very beginning of one autistic participant's story:

> Interlocutor 1: . . .I've actually never seen it before because we had another RA pick out these videos. So, what happened?

Participant: Oh yeah. There was the light and the ball...

Since ambiguous references like this one occurred at the very beginning of the narrative, it is unlikely they are caused by the participant's inability to keep track of referential mentions in the preceding discourse. Relatedly, if referential ambiguity by autistic storytellers is due to relative weaknesses in executive functioning skills, we should have seen relatively high rates of competing-referents ambiguity in the autistic group. As is clear in the following example, competing-referents ambiguity is more likely later in the discourse, once multiple characters have been established.

Interlocutor 2: ...What was the video you just watched about?

Participant: Oh yeah, the second one was about, like, [the light]$_j$...

[[The little light]$_j$ and [the big light]$_k$]$_i$ were [companions]$_i$ because [they]$_i$ were playing catch ...

[They]$_i$ were using [the beach ball]$_l$ which was easier, but then when [the little lamp]$_j$ tried to hit [the big green bowling ball]$_m$ [it]$_{j/m}$ pushed [the lamp]$_{j/k}$ out...

Group differences in executive functioning would predict more challenges with competing-referents ambiguity in the autism group, as participants must track multiple characters and previous references to them. The idea that executive functioning weaknesses do contribute to competing-referents ambiguity is supported by the findings from the work of Kuijper et al. (2015), where children with ADHD were more likely to produce competing-referents ambiguity. However, these authors, like us, did not find more frequent uses of competing-referents ambiguity by autistic participants compared to NA participants. More recently, Stegenwallner-Schütz and Adani (2020) used a picture description task to probe for competing-referents ambiguity, specifically, and also found that autistic children were not more likely to produce competing-referents ambiguity than NA children. These authors also established that working memory capacity significantly predicts the likelihood of producing competing-referents ambiguity in both groups, again suggesting that competing-referents ambiguity is associated with executive functioning. Therefore, the fact that this study and the studies of Kuijper et al. (2015) and Stegenwallner-Schütz and Adani (2020) all found that this type of ambiguity was not particularly prevalent in autistic narratives calls into question whether executive functioning differences can explain increased referential ambiguity in autistic narratives.

Finally, differences in rates of ambiguity subtypes have practical implications for both clinicians and researchers. Although we present our findings with a focus on differences between autistic and NA storytellers' use of references, a larger takeaway is that never-introduced–referent and competing-referents ambiguity are distinct subtypes of referential ambiguity that can be differentially represented in individuals' narratives. This differentiation implies that each may depend on a separate underlying ability: considering the background knowledge of a given listener and tracking reference in one's own narrative, respectively. Therefore, clinicians and researchers who are assessing referential term production, for any population, should score each subtype of ambiguity independently and should ensure that their testing procedure/protocol can sufficiently and accurately probe for both types. Regarding the latter point, we encourage the elicitation of stories involving several characters who each first appear at varying points in the narrative. This will allow for ample opportunities for competing referents and never-introduced referents, respectively. We also warn against using story generation tasks that involve visual prompts, especially when those prompts are visible to the examiner, as this could encourage the use of deictic references (e.g., saying "the man" while pointing or looking at a picture).

We are more cautious about offering intervention recommendations because stories in this study were not assessed for clarity or quality. It is possible that perceptions of narrative quantity/quality are not attributable to the relative frequency of either type of ambiguity, but instead to particular ambiguous references—ones that are crucial to the specific narrative context. We discuss this more in our Limitations section.

Lexical Categories

The focus of many previous studies has centered specifically on differences in pronoun use by autistic individuals, but our findings show that referential ambiguity in the narratives produced by autistic participants was not fundamentally dependent on ambiguous use of pronouns. In fact, percentages of ambiguous pronouns were similar between groups (approximately 28% in the ASD group and 23% in the NA group), and—importantly—the proportion of pronouns used that were ambiguous due to their having competing referents was very similar (20% ASD vs. 21% NA). Across both participant groups, pronouns were significantly more ambiguous due to having competing referents, which was not the type of ambiguity that was relatively frequent in the narratives produced by autistic children.

Thus, at face value, our lexical category findings seem to reinforce existing findings: Autistic participants were significantly more likely to produce never-introduced–referent ambiguity, and because pronouns are less likely to be used this way, autistic children did not overproduce ambiguous pronouns. However, the lexical category findings also introduce some insight into the nature of referential differences in stories told by autistic narrators. Because we find that never-introduced ambiguous pronouns were significantly less frequent than competing-referents ambiguous pronouns in both groups, we show that autistic participants were not likely to use pronouns to refer to never-introduced characters. This replicates previous findings, where studies show that autistic children are unlikely to use pronouns to introduce characters for the first time (Arnold et al., 2009; Kuijper et al., 2015). Similarly, a recent article measuring pronoun comprehension/processing in autism in discourse contexts (Nagano et al., 2021) finds that older children with ASD and age/IQ/language-matched NA peers rely on similar principles to guide their interpretation of pronouns in discourse. Thus, it is unlikely that autistic individuals have a different or weaker understanding of the principles guiding anaphora selection. Instead, autistic participants in this study seemed to have more difficulty with definiteness, so that they tend to use a definite NP or a proper noun in contexts that do not allow for them:

> Interlocutor 2: Hi, sorry for the delay. What have you guys been up to?
>
> Participant: I watched, like, the two videos.
>
> Interlocutor 2: The two videos? Oh! You must mean the videos that the student put on the desktop of that computer. I haven't seen them. . .What was the video you just watched about?
>
> Participant: Oh yeah, the second one was about, like, the light. . .The little light and the big light were companions because they were playing catch. . .They were using the beach ball which was easier, but then when the little lamp tried to hit the big green bowling ball it pushed the lamp out. . .

This participant uses definite NPs from the very beginning of their story (all inappropriate uses of definite determiners are underlined), including the use of a definite NP to reference the videos in the first place, even though it had earlier been made clear to the participant that no one currently working in the lab was familiar with the videos being shown and even though the current interlocutor (Interlocutor 2) had not been in the room at the initiation of either video. This finding replicates a pattern described by Norbury and Bishop (2003) and Tager-Flusberg (1995), in which a relatively large proportion of the autistic individuals in their studies used definite NPs to introduce a character.

Most never-introduced ambiguity was due to the inappropriate use of definite determiners, but a few participants also referred to the characters (or videos) by a proper name, even though these names were not familiar to the researcher. Consider the following example:

> Interlocutor 1: How did you like that video? I've never seen it before.
>
> Participant: That was Luxo Jr., right?

By using definite nouns and proper names at the beginning of their stories, autistic participants show a pattern of being overspecific, by referencing a specific entity with whom their listener is not yet familiar, rather than being underspecific, by using a pronoun. This finding corresponds with results from Arnold et al. (2009) and Baltaxe (1977), who both find that autistic participants overspecify referents by using full NPs when a pronoun would be appropriate. Neither of these previous works analyze definiteness or ambiguity, so the addition of our findings to theirs suggests that autistic storytellers may be sometimes ambiguous because they tend to overspecify, by using full, definite NPs in cases where an indefinite NP would be more effective at establishing reference. Although the idea of ambiguously specific reference may seem oxymoronic, specificity can in fact lead to ambiguity: If a speaker uses a too specific NP, such as a definite NP or a proper noun, in cases where their listener cannot possibly access the intended referent, the result is ambiguity.

Again, findings have some practical implications. Ever since Kanner (1943) described differences in pronoun production by autistic children, clinicians and researchers have been paying special attention to pronoun use by this population. However, the findings of the current work suggest that, for older autistic children, referential ambiguity is driven not by the misuse of pronouns but by differences in their use of definite NPs, including proper names. We encourage future research to build on this finding, by examining how autistic individuals use/understand rules that guide definiteness, and we suggest that clinicians consider assessing/treating these rules as well.

Narrative Sequence

A previous study showed that autistic speakers tend to struggle with the task of updating referential terms to

meet the need of a new listener, after having previously established reference with someone else (Nadig et al., 2015). In contrast, our narrative relay paradigm yielded similar rates of ambiguity across both narratives. In fact, the effect of narrative sequence (and therefore the effect of switching listeners) had no statistically significant consequences at all, except that all participants used more referential terms in the second narrative. This was likely caused by the relative complexity of the second video, which involved both lamps interacting with two balls at the same time. In the first video, a second ball is only introduced at the very end, after a previous one has deflated, so there were maximally three entities to describe and track at once. Because autistic participants tended to initially introduce characters with definite determiners and proper nouns for both narratives and because the current paradigm involved two listeners who were both introduced as completely naïve to the content of the videos being described, the likelihood of never-introduced–referent ambiguity was equally possible across both listeners. We suspect that results would be different if at least one of the listeners was not explicitly declared to be naïve.

Limitations and Future Directions

A primary limitation of the current work is that it is based on a small number of participants, who are characterized by a very specific linguistic and cognitive profile. We instituted strict participant selection criteria to try to ensure that group differences could not be attributable to other factors (e.g., general language ability), but this resulted in a relatively small pool of participants, whose characteristics are not representative of either the autistic or NA population, generally. Furthermore, effect sizes were moderate. Thus, current findings cannot and should not be interpreted as representing the "final answer" with regard to narrative reference production in autism and should not be assumed to generalize to all autistic or NA people.

Other limitations pertain to study procedures, including coding and subsequent analyses. First, in this study, the two listeners were similar to one another: They were both NA, and they were both presented as naïve to the videos. This meant that our narrative relay condition could not tell us much about participants' abilities to adjust referential strategies to meet the needs of different types of listeners. Future work could investigate the effect of between-listener variability, such as systematically varying the background knowledge of interlocutors and/or pairing both autistic and NA storytellers with both autistic and NA listeners. Second, researchers did not ask or remind participants to be as clear as possible, nor did researchers display confusion or ask clarifying questions (e.g., "Wait, who pushed the ball?"). We did the first to maintain a naturalistic context, as it is unlikely that listeners in "real life" would preface their request to hear a story with the reminder that their narrator be clear. We did the latter to maintain consistent narrative elicitation procedures across participants. Avoiding requests for clarification also made coding and analysis more straightforward because we avoided the added complication of coding/analyzing whether a given referent was being produced before or after a listener expressed confusion or asked for clarification. It is possible that both manipulations would affect referential ambiguity. Perhaps rates of ambiguity—overall or for one of the two ambiguity subtypes—would decrease when speakers are initially reminded to be clear or after the listener demonstrates confusion. Such findings would give insight into whether speakers are able to avoid or repair certain types of ambiguity. To address this, future researchers should consider comparing ambiguity rates between conditions where storytellers receive cues regarding clarity to those when they do not. They could also (or instead) tell participants ahead of time that their stories would be critiqued or judged. This, again, may drive down rates of ambiguity. Such findings would show that participants have knowledge about how to make references unambiguous, and this would suggest that current findings reflect participants simply forgetting or neglecting to do so in the context of this study. Finally, we asked participants to provide two tokens of a very specific type of story: a story about fictional characters involved in a series of somewhat predictable events (e.g., they play with an inflatable ball, and it deflates). Ambiguity rates could differ across narrative types. For example, personal narratives (when told to a stranger) involve no expectation of listener knowledge of either events or characters. In contrast, retellings of popular stories (e.g., Little Red Riding Hood) involve the opposite expectation. These methodological changes, and others, could change the underlying processes by which people determine how to use referential language and could therefore result in different ambiguity rates.

Another limitation of this study is that we did not measure executive functioning or perspective taking. Both of these metrics could correlate with the two types of ambiguity we analyzed. Therefore, although we suggest that patterns of ambiguity in narratives told by this set of autistic participants imply differences in attending to listeners' needs versus challenges in tracking reference, we cannot fully confirm this without measuring these cognitive abilities on their own. We encourage future researchers to address this by testing our proposed relationship between cognitive skills and the two subtypes of referential ambiguity.

Finally, we encourage future research to explore the relationship between referential ambiguity and story clarity and quality, perhaps by asking listeners to rate stories,

answer comprehension questions, and/or retell story events. While the current project provides insight into relative rates of ambiguity subtypes between groups, it does not show how these rates contribute to a listener's perception of story goodness or to their understanding of characters and events. Although we assume that higher rates of ambiguity would, in general, predict higher rates of listener confusion, it is possible that listener understanding could be negatively affected by even a single ambiguous referent, if that referent is central to story events. In other words, perceptions of quality and comprehension accuracy may be driven by the quality of ambiguity, rather than the quantity.

Conclusions

Current findings offer several important contributions to our understanding of referencing in narratives told by autistic storytellers. First, when tasked with telling a story that contains multiple characters that can be referenced using the same phrases, autistic narrators are no more ambiguous than NA narrators; this suggests that autistic narrators are just as competent as NA ones at maintaining referential clarity throughout their narratives, at least in terms of rates of ambiguity. Instead, differences in referential clarity between groups are not due to the overall frequency of ambiguity but the type of ambiguity. Specifically, autistic narrators use significantly more referential phrase forms (e.g., definite NPs) in contexts when a character has not yet been introduced, leading to never-introduced ambiguity. In contrast, autistic narrators are as successful as NA speakers at clarifying reference when there are competing antecedents available for a given anaphoric NP, suggesting similar levels of proficiency in this skill between groups. The important difference between these two types of ambiguity is that the first is listener/context specific, whereas the second is generic. This finding has practical implications as well: Clinicians (and researchers) who are assessing and targeting in treatment children's narrative skills should make sure to separately consider never-introduced–referent and competing-referents ambiguity, as each likely stems from different underlying challenges and as some populations may produce increased instances of one type, but not both. Finally, our work indicates that referencing differences between autistic and NA children are not limited to, or even dominated by, differences in the use of pronouns. Instead, our work highlights differences in the use of all NPs and especially in the use of definiteness. These data provide important insights into the nuances of referential use by autistic individuals and highlight the importance of investigating communicative abilities across multiple pragmatic and linguistic contexts.

Data Availability Statement

The data sets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Acknowledgments

Research was funded by National Institute on Deafness and Other Communication Disorders Grant 1R01DC012774-01 to Ruth B. Grossman. The authors are incredibly grateful to the children and families who participated in this research. The authors would also like to thank Riley Myhaver, Kimberly Clark, Lindsey Filbey, and Anna Privett for transcription and coding.

References

Abbott, B. (2004). Definiteness and indefiniteness. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 122–150). Blackwell.

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review, 106*(4), 748–765. https://doi.org/10.1037/0033-295X.106.4.748

Andrés-Roqueta, C., & Katsos, N. (2020). A distinction between linguistic and social pragmatics helps the precise characterization of pragmatic challenges in children with autism spectrum disorders and developmental language disorder. *Journal of Speech, Language, and Hearing Research, 63*(5), 1494–1508. https://doi.org/10.1044/2020_JSLHR-19-00263

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

Baltaxe, C. A. M. (1977). Pragmatic deficits in the language of autistic adolescents. *Journal of Pediatric Psychology, 2*(4), 176–180. https://doi.org/10.1093/jpepsy/2.4.176

Baltaxe, C. A. M., & D'Angiola, N. (1996). Referencing skills in children with autism and specific language impairment. *International Journal of Language & Communication Disorders, 31*(3), 245–258. https://doi.org/10.3109/13682829609033156

Baltaxe, C. A. M., Russell, A., D'Angiola, N., & Simmons, J. Q. (1995). Discourse cohesion in the verbal interactions of individuals diagnosed with autistic disorder or schizotypal personality disorder. *Australia and New Zealand Journal of Developmental Disabilities, 20*(2), 79–96. https://doi.org/10.1080/07263869500035471

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

Beechey, T. (2022). On perspective taking in conversation and in research: A comment on Bambara et al. (2021). *Journal of Speech, Language, and Hearing Research, 65*(4), 1597–1599. https://doi.org/10.1044/2021_JSLHR-21-00606

Brugman, H., & Russel, A. (2004). *Annotating multimedia/multimodal resources with ELAN*. Proceedings of LREC 2004,

Fourth International Conference on Language Resources and Evaluation.

Colle, L., Baron-Cohen, S., Wheelwright, S., & Van Der Lely, H. K. J. (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

ELAN. (2022). *Nijmegen: Max Planck Institute for Psycholinguistics*. The Language Archive. https://archive.mpi.nl/tla/elan

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

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test: KBIT 2*. Pearson.

Kuijper, S. J., Hartman, C. A., & Hendriks, P. (2015). Who is he? Children with ASD and ADHD take the listener into account in their production of ambiguous pronouns. *PLOS ONE, 10*(7), Article e0132408. https://doi.org/10.1371/journal.pone.0132408

Lasseter, J. (1986). *Luxo Jr*. Pixar. https://www.youtube.com/watch?v=6G3O60o5U7w

Lasseter, J., & Stanton, A. (1991). *Light and Heavy*. Pixar Animation Studios. https://www.youtube.com/watch?v=M4NmpDv_GYE

Lord, C., DiLavore, P. C., Gotham, K., Guthrie, W., Luyster, R. J., Risi, S., & Rutter, M. (2012). *Autism Diagnostic Observation Schedule: ADOS-2*. Western Psychological Services.

Malkin, L., Abbot-Smith, K., & Williams, D. (2018). Is verbal reference impaired in autism spectrum disorder? A systematic review. *Autism & Developmental Language Impairments, 3*. https://doi.org/10.1177/2396941518763166

Mäkinen, L., Loukusa, S., Leinonen, E., Moilanen, I., Ebeling, H., & Kunnari, S. (2014). Characteristics of narrative language in autism spectrum disorder: Evidence from the Finnish. *Research in Autism Spectrum Disorders, 8*(8), 987–996. https://doi.org/10.1016/j.rasd.2014.05.001

Meyer, M. (1969). *Frog, where are you?* Dial Press.

Milton, D. E. (2012). On the ontological status of autism: The "double empathy problem." *Disability & Society, 27*(6), 883–887. https://doi.org/10.1080/09687599.2012.710008

Nadig, A., Seth, S., & Sasson, M. (2015). Global similarities and multifaceted differences in the production of partner-specific referential pacts by adults with autism spectrum disorders. *Frontiers in Psychology, 6,* Article 1888. https://doi.org/10.3389/fpsyg.2015.01888

Nagano, M., Zane, E., & Grossman, R. B. (2021). Structural and contextual cues in third-person pronoun interpretation by children with autism spectrum disorder and their neurotypical peers. *Journal of Autism and Developmental Disorders, 51*(5), 1562–1583. https://doi.org/10.1007/s10803-020-04645-7

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/10.1080/1368203100001081 33

Novogrodsky, R. (2013). Subject pronoun use by children with autism spectrum disorders (ASD). *Clinical Linguistics & Phonetics, 27*(2), 85–93. https://doi.org/10.3109/02699206.2012.742567

Novogrodsky, R., & Edelson, L. R. (2016). Ambiguous pronoun use in narratives of children with autism spectrum disorders. *Child Language Teaching and Therapy, 32*(2), 241–252. https://doi.org/10.1177/0265659015602935

RStudio Team. (2022). *RStudio: Integrated Development Environment for R (2022.2.3.492)*. RStudio, PBC. http://www.rstudio.com/

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire (SCQ)*. Western Psychological Services.

Stegenwallner-Schütz, M., & Adani, F. (2020). Production of referring expressions by children with ASD: Effects of referent accessibility and working memory capacity. *Language Acquisition, 27*(3), 276–305. https://doi.org/10.1080/10489223.2020.1769625

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

Tager-Flusberg, H. (1995). 'Once upon a ribbit': Stories narrated by autistic children. *British Journal of Developmental Psychology, 13*(1), 45–59. https://doi.org/10.1111/j.2044-835x.1995.tb00663.x

Volden, J., Mulcahy, R. F., & Holdgrafer, G. (1997). Pragmatic language disorder and perspective taking in autistic speakers. *Applied Psycholinguistics, 18*(2), 181–198. https://doi.org/10.1017/S0142716400009966

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

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals—Fifth Edition (CELF-5)* (5th ed.). Pearson.