*Applied Psycholinguistics* (2026), **47**, e30, 1–31
doi:10.1017/S014271642610068X

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

# Online pronoun resolution in children with developmental language disorder (DLD): A visual world eye-tracking study

Llorenç Andreu[1], Aurora Bel[2], Ernesto Guerra[3,4], Nadia Ahufinger[1,5] and Monica Sanz-Torrent[6]

[1]eHealth Centre: Human and Planetary Health Research Centre, Universitat Oberta de Catalunya, Barcelona, Spain, [2]Departament de Traducció i Ciències del Llenguatge, Universitat Pompeu Fabra, Barcelona, Spain, [3]Center for Advanced Research in Education, Institute of Education (IE), Universidad de Chile, Santiago, Chile, [4]Departamento de Lingüística, Facultad de Filosofía y Humanidades, Universidad de Chile, Santiago, Chile, [5]Estudis de Psicologia i Ciències de l'Educació, Universitat Oberta de Catalunya, Barcelona, Spain and [6]Departament de Cognició, Desenvolupament i Psicologia de l'Educació, Secció Cognició, Universitat de Barcelona, Barcelona, Spain
**Corresponding author:** Llorenç Andreu; Email: landreub@uoc.edu

(Received 26 October 2025 UTC; revised 12 March 2026 UTC; accepted 8 May 2026 UTC)

**Abstract**

Children with Developmental Language Disorder (DLD) often experience difficulties with morphosyntax and discourse processing, which hinder their ability to establish referential coherence. While pronoun resolution has been extensively studied in typically developing (TD) children, little is known about how children with DLD process pronouns in real time, especially in Spanish—a language with rich morphology and flexible word order. This study investigated how Spanish-speaking children with DLD interpret third-person subject pronouns during sentence comprehension, examining their use of semantic and syntactic cues in reference resolution. Across three eye-tracking experiments, we tested children's reliance on semantic gender cues in overt pronouns (Experiment 1), on syntactic cues such as grammatical role and order of mention in overt pronouns (Experiment 2), and on these same cues in null pronouns (Experiment 3). Participants were 48 Spanish–Catalan bilingual children: 16 with DLD, 16 age-matched TD peers, and 16 younger TD children matched by mean length of utterance. Eye movements during a visual world task were analyzed using growth curve and mixed-effects models. Results showed that children with DLD used semantic gender cues when available and relied on first-mentioned or subject referents when such cues were absent, suggesting compensatory use of structural heuristics in pronoun resolution.

**Keywords:** Developmental language disorder (DLD); eye-tracking; pronoun resolution; Spanish

## Introduction

Developmental Language Disorder (DLD), previously referred to as Specific Language Impairment (SLI), is a neurodevelopmental condition that affects

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

approximately 7.5% of children (Norbury et al., 2016; Tomblin et al., 1997). Recent studies suggest that DLD affects boys and girls at similar rates (Calder et al., 2022; Norbury et al., 2016). It is characterized by significant and lasting difficulties in acquiring and developing oral language that cannot be explained by other medical conditions such as hearing impairment, intellectual disability, autism, neurological disorders, or genetic syndromes (Bishop et al., 2016). Additionally, DLD frequently co-occurs with other neurodevelopmental disorders, including attentional problems (ADHD), motor problems (developmental coordination disorder or DCD), reading and spelling problems (developmental dyslexia), speech problems, limitations of adaptive behavior, and/or behavioral, and emotional disorder (Bishop et al., 2017).

Children diagnosed with DLD frequently face significant challenges in both understanding and producing spoken language. These difficulties span various domains of language and cognition, including phonology, morphology, syntax, semantics, vocabulary, pragmatics, discourse, and verbal learning and memory (Bishop et al., 2017; Leonard, 2014). Extensive research has examined morphological difficulties in children diagnosed with DLD. These children often have trouble using inflectional morphology correctly, particularly with verb tense and subject-verb agreement (e.g. Abel et al., 2015; Van der Lely & Ullman, 2001). Moreover, several studies have also found that children with DLD frequently leave out functional words such as articles, prepositions, and pronouns (e.g. Bedore & Leonard, 2001; Coloma et al., 2016; Restrepo & Gutierrez-Clellen, 2001; Sanz-Torrent et al., 2008). In relation to pronouns, several cross-linguistic studies have shown that children with DLD struggle with the production of different kinds of pronouns. Previous research has documented a higher incidence of errors in pronoun case marking for third-person singular pronouns (e.g., *him runs* vs. *he runs*) in English-speaking children with DLD compared to their same-age peers (e.g., Loeb and Leonard, 1991; Moore, 1995, 2001). In French children with DLD, Jakubowicz et al. (1998) found that they produced significantly fewer clitics in object position than clitics in subject position and reflexive pronouns. Additionally, Paradis et al. (2006) replicated these difficulties in the domain of object clitics in French–English bilingual children with DLD. The same difficulties have been found in Italian (Bortolini et al., 2006; Cipriani et al., 1998) and Greek children with DLD (Stavrakaki, 1999; Stavrakaki & Van der Lely, 2010). In Spanish, a null-subject language in which personal pronouns can appear either explicitly or as null forms, children with DLD produce a higher rate of null subject expression from spontaneous production than TD children (Grinstead et al., 2018). Additionally, children with DLD produce fewer overt subject pronouns in switch-reference contexts—situations where the subject of the preceding finite verb differs from the subject of the current clause—compared to TD peers (Dickinson et al., 2023).

However, while numerous studies have investigated different aspects of pronoun production development in children with DLD, relatively little attention has been given to pronoun resolution (Abbott et al., 2024; Rakhlin et al., 2015; van der Lely & Stollwerck, 1997). Pronoun resolution is a key component of language comprehension and discourse processing. It involves identifying the antecedent—the word or phrase to which a pronoun relates (anaphor). This process enables listeners and readers to maintain coherence across sentences and larger discourse units. In this

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

sense, pronouns pose a unique challenge because they are inherently underspecified and rely heavily on contextual cues for interpretation.

Previous studies have found that children with DLD have difficulties with correct pronoun reference assignment. Van der Lely and Stollwerck (1997) investigated the intrasentential assignment of reference to pronouns (*him, her*) and reflexives (*himself,* herself) in 9–12-year-old English-speaking children with DLD in a picture–sentence pair judgment task. Participants were presented with pictures containing two characters (e.g., Mowgli and Baloo Bear). Then, an introductory sentence was spoken by the experimenter (e.g., "This is Mowgli; this is Baloo Bear"), which was followed by the experimental test sentence (e.g., "Is Mowgli tickling him?") to which the children had to reply yes/no as appropriate. Children with DLD had difficulty identifying the correct antecedent for pronouns and reflexives when relying solely on syntactic cues. However, their performance improved notably when they could rely on semantic gender cues based on sex distinctions to identify the referent, particularly in scenarios where the introductory sentence presents two characters of different sexes (e.g., "This is Peter Pan; this is Wendy. Is Peter Pan touching her?"). This finding apparently contrasts with results from other studies showing that children with DLD often struggle with grammatical gender and other morphosyntactic cues. For instance, Bedore and Leonard (2001) reported such difficulties in Spanish-speaking children with DLD, highlighting persistent challenges in using gender agreement and other grammatical markers. Taken together, these observations suggest that while children with DLD can effectively draw on semantic gender information—where gender distinctions map onto real-world categories such as biological sex—their ability to rely on grammatical gender, which is more arbitrary and morphosyntactic in nature, remains inconsistent and limited.

Similarly, Rakhlin et al. (2015) examined intrasentential reference assignment to pronouns and reflexives in 5–10-year-old native-Russian-speaking children with DLD on a picture selection task. The study revealed that children with DLD performed worse overall than their TD peers, with pronouns and reflexives being relatively similar in difficulty for the DLD, but not the TD group which showed more difficulties in pronouns than reflexives. Additionally, the challenges the children faced were not only due to the existence of anaphoric relationships, but also, in part, to the increased syntactic complexity of the sentences involved (e.g., "The boy patted him" vs. "The boy sees him dancing."). More recently, Abbott et al. (2024) examined pronominal resolution during real-time sentence processing and overt comprehension and whether the rate of speech impacted access to that knowledge in 6–11-year-old English-speaking children with DLD in two experiments. In the first experiment, participants performed a binary picture-matching task during uninterrupted auditory sentence processing, with response times (RTs) measured at the onset of pronouns, reflexives, or control nouns. To assess facilitative priming, RTs in the pronoun condition were compared against the control noun baseline across two speech rates. While children with DLD exhibited no evidence of real-time pronoun-antecedent priming at a regular speech rate, they demonstrated significant facilitative priming when the auditory stimuli were slowed. The second experiment used a sentence-picture matching task to assess comprehension accuracy of similar sentences containing a pronoun. Children with DLD performed at chance levels for pronoun comprehension, regardless of the speech rate. Therefore,

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

slowing the speech helped children with DLD with real-time processing, but not with overall comprehension accuracy.

Although some studies have explored pronoun resolution in children with DLD, there is still a lack of systematic research examining how these children use different linguistic and cognitive cues to resolve pronoun dependencies. In contrast, linguistic and psycholinguistic research has extensively examined the mechanisms guiding pronoun interpretation in both TD children and adults, highlighting the interaction of structural, cognitive, and pragmatic factors. Children begin producing pronouns early in development, but accurate interpretation emerges more gradually. Initial strategies show a robust preference for linking pronouns to the subject or first-mentioned entity, often referred to as a prominence heuristic or agent bias (MacWhinney & Bates, 1989; Song & Fisher, 2005), even when the discourse-preferred antecedent is not the subject. Eye-tracking studies reveal that children aged 4–5 do not yet integrate syntactic and pragmatic cues in an adult-like way. For example, Arnold et al. (2007) found that children did not exhibit anticipatory fixations toward likely referents and instead relied heavily on gender congruence, indicating a dependence on surface-level features rather than deeper structural or discourse information. In null-subject languages like Spanish, children exhibit early sensitivity to language-specific patterns: null pronouns are typically interpreted as referring to subjects, while overt pronouns (e.g., *él, ella*) tend to be linked with non-subjects (Filiaci & Baldo, 2009; Grüter, 2006). This reflects early awareness of pragmatic conventions, where null pronouns maintain topic continuity and overt pronouns often mark a topic shift. However, the acquisition of this referential system is protracted due to the low frequency of overt pronouns in child-directed input, limiting opportunities for probabilistic learning. Longitudinal studies in two very close languages, such as Spanish and Catalan (Bel et al., 2010; Bel & Albert, 2016), confirm that children do not consistently use structural and discourse cues to establish coreference in extended narratives until around 10–12 years of age, mirroring findings from other languages (Hickmann, 2004; Serratrice, 2007; Shin & Cairns, 2012). This protracted development reflects the multifaceted nature of referential interpretation, which necessitates the integration of morphosyntactic, pragmatic, and cognitive resources, such as Theory of Mind (Hendriks et al., 2014).

In adults, pronoun interpretation is characterized by a more refined weighting and rapid integration of the same grammatical, cognitive, and pragmatic cues that children are still learning to coordinate. Features such as gender and number contribute to identifying antecedents (Carreiras et al., 1993; Crawley et al., 1990), and visual-world eye-tracking studies show that adults make anticipatory fixations based not only on gender cues (Arnold et al., 2000) but also on morphosyntactic structure. Two well-established strategies—subject preference and first-mention bias—play key roles: the former reflects the syntactic prominence of subjects, while the latter captures a cognitive tendency to link pronouns to the first-mentioned entity (Bader & Portele, 2019; McDonald & MacWhinney, 1995; Stevenson et al., 1994). Although these strategies often align, they are driven by distinct mechanisms—structural versus linear order. Studies in languages with flexible word order have helped disentangle these influences. In Finnish, subject preference tends to emerge earlier (Järvikivi et al., 2005; Kaiser & Trueswell, 2003), while recent findings in Catalan (Bel et al., 2026) suggest a stronger role for order of mention.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

Pragmatic factors such as topic and focus also shape referent prominence. Sentence structure—including canonical versus non-canonical forms and constructions like clefts—affects how antecedents are processed. While some studies suggest that focus enhances referent salience (Káldi & Babarczy, 2021), others indicate that increased prominence does not automatically translate into a referential preference; for instance, clefting can actually reduce the likelihood of a referent being linked to a pronoun, producing an "anti-focus effect" (Colonna et al., 2015; de la Fuente, 2015; Hert et al., 2024; Patterson et al., 2017). In Spanish, de la Fuente (2015) found that null pronouns are more likely to refer to non-clefted antecedents, underscoring the role of pragmatic structure. However, findings on the effect of clefts are mixed, with some studies reporting no significant impact (Colonna et al., 2015; Järvikivi et al., 2014). Altogether, these studies indicate that adults rely on a nuanced integration of morphosyntactic, semantic, and pragmatic cues to resolve pronouns efficiently—a capacity that contrasts with the slower, more limited strategies observed in children.

The present study examines how children with DLD interpret pronouns in real time using eye-tracking within the Visual World Paradigm (VWP; Tanenhaus et al., 1995). Real-time eye-tracking measures of spoken language processing can provide valuable insights into how children with DLD resolve referential ambiguity. This method enables the capture of moment-by-moment gaze patterns as children listen to sentences that describe a visual scene (Trueswell, 2008). In comparison with previous offline studies that used picture-selection tasks (Rakhlin et al., 2015; Van der Lely & Stollwerck, 1997) and the earlier online study employing a cross-modal picture priming task (Abbott et al., 2024), the VWP offers clear advantages. It integrates the presentation of spoken input with a near-continuous, naturalistic measure of comprehension—namely, participants' eye movements toward relevant referents. Furthermore, because the task does not require explicit responses such as pointing or button pressing, it is particularly well-suited for studying children with DLD. As noted by Schul et al. (2004), these children often exhibit slower visual–motor responses but preserved visual–attentional shifting, indicating that eye-tracking can reveal online language processing mechanisms without the confounding effects of motor demands.

## Research questions and hypotheses

The present study investigates the real-time processing of third-person personal pronouns in Spanish-speaking children with DLD. We focus specifically on subject-position pronouns, which in Spanish can be realized as either overt (*él/ella*) or null (ø) forms. We address the following research questions and hypotheses across three experiments:

*Research Question 1 (Experiment 1)*

To what extent do children with DLD utilize local semantic cues, specifically grammatical gender, to resolve pronoun dependencies in unambiguous contexts?

- *Hypothesis 1*: Based on findings that children with DLD can leverage salient surface features (Arnold et al., 2007; Van der Lely & Stollwerck, 1997), we hypothesize that they will successfully use gender cues to assign referents when the two potential antecedents differ in gender.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

*Research Question 2 (Experiment 2)*

How do children with DLD utilize structural and linear cues, specifically subjecthood and order of mention, when interpreting overt pronouns?

- *Hypothesis 2*: Drawing on evidence of reduced referential accuracy in DLD populations (Rakhlin et al., 2015), we hypothesize that children with DLD will struggle to integrate these abstract cues. Specifically, we predict they will not exhibit a consistent first-mention bias or subject-preference when resolving overt pronouns, showing less sensitivity to discourse prominence than TD peers.

*Research Question 3 (Experiment 3)*

Do children with DLD demonstrate sensitivity to language-specific pragmatic conventions when resolving null pronouns?

- *Hypothesis 3*: Given the protracted development of the null-subject system and the multifaceted nature of topic continuity, we hypothesize that children with DLD will show delayed or inconsistent use of structural and discourse cues (such as the subject-bias for null forms) compared to the more robust patterns expected in TD children.

## Experiment 1

### *Method*

*Ethics statement and data availability*

The research protocol was approved by the Ethics Committee of the Universitat Oberta de Catalunya (UOC). Parents or legal guardians of all participating children provided written informed consent for both data collection and analyses. The materials, data, and scripts used in this study are available on the Open Science Framework (OSF) at the following link: https://osf.io/gh68q/.

*Participants*

Participants included a sample of 48 children (16 girls and 32 boys) that was divided into three groups: 16 children with DLD (5 girls and 11 boys; age range = 4;07–12;03 years, average = 7 years and 8 months; hereafter "DLD group"), 16 typically developing age-matched children (6 girls and 10 boys; age range = 4;02–12;00 years, average = 7 years and 8 months; hereafter "Age group"), and 16 younger children matched by mean length of utterance by words (MLU-w) and sex with the DLD group (5 girls and 11 boys; age range = 4;00–7;06 years, average = 5 years and 7 months; hereafter "MLU group"). All children resided in the metropolitan area of Barcelona, where both Catalan and Spanish are spoken. In Catalonia, both Catalan and Spanish are recognized as official languages, and children generally achieve native or near-native proficiency in each. Consequently, truly monolingual speakers are rare, except among recent arrivals. Bilingualism in this setting is typically balanced, as most people display comparable competence in both languages relative to monolingual peers of the same age.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

Catalan serves as the main medium of instruction in schools, with Spanish also taught as part of the curriculum. Despite Catalan's strong institutional promotion and official status, Spanish continues to play a dominant role in everyday communication. Both languages are used freely across social contexts. Participants met the following inclusion criteria: (1) normal hearing as reported by parents; (2) normal or corrected-to-normal vision as reported by parents; (3) normal oral and speech motor abilities; and (4) native bilingual Catalan–Spanish speakers as reported by parents. Information about each child's bilingual background and language use was collected through a questionnaire completed by parents. They indicated the main language used by both the mother and the father when interacting with the child, as well as the child's dominant language and the total number of languages spoken at home. All families reported that their children were able to both speak and understand Catalan and Spanish. This allowed researchers to obtain a comprehensive overview of each child's linguistic exposure within the home environment. Moreover, before the formal assessment, research assistants carried out a short preliminary task involving around ten minutes of spontaneous conversation with each child. Using visual prompts related to familiar themes—such as family life, school activities, and personal interests—this procedure facilitated the collection of natural speech samples and enabled the evaluation of comprehension and expressive language in both Catalan and Spanish. This assessment served as an inclusion criterion; while several potential candidates were excluded during this initial screening for not meeting proficiency requirements in both languages, only the 48 children who demonstrated adequate understanding and verbal production in both Catalan and Spanish were included in the final sample reported here. In addition, Children with biomedical conditions linked to genetic or neurological causes—such as autism, intellectual disability, or Down or Williams syndromes—were excluded from the study.

All participants were assessed with cognitive and language standardized tests: The Kaufman Brief Intelligence Test (KBIT, Spanish version; Kaufman & Kaufman 2004), the Peabody Picture Vocabulary Test Spanish version—Third Edition (PPVT-III; Dunn et al. 2006), and the Comprehension Test of Grammatical Structures (CEG; Mendoza et al., 2006) that is a Spanish adaptation of the TROG—Test for Reception of Grammar (Bishop, 2003). Additionally, grammatical expressive complexity was assessed by calculating the MLU-w (mean length of utterance in words) for each child based on their spontaneous speech productions.

### *DLD group status confirmation*

Children in the DLD group included those with a prior diagnosis of DLD as well as those with language difficulties referred by speech-language therapists from schools and hospitals. Participants were recruited from three distinct institutions: (1) UTAE (Learning Disorders Unit at Sant Joan de Déu-Barcelona Children's Hospital) that is a specialized unit that provides evaluation, diagnosis, and personalized interventions for children and adolescents with language, cognitive and behavioral difficulties that affect learning; (2) CREDA (Centers of Resources for Language-and Hearing-Impaired Children) that are public teams of speech and language pathologists that collaborate with schools across Catalonia to assess, identify and provide

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

intensive intervention to the children with DLD; and (3) ATELCA (Catalan Association for Developmental Language Disorder), an association of families with children who have DLD.

The criteria to confirm the participants' DLD status were based on multiple data sources and were confirmed by two trained investigators specialized in child language development. First, children in the DLD group had to score a non-verbal intellectual quotient (NVIQ) ≥ 70 (Nonverbal Kaufman Brief Intelligence Test [K-BIT; Kaufman & Kaufman, 2004]) following Bishop et al. (2016). Second, for language status, they had to demonstrate oral language difficulties based on scores at least 1.25 SD below the mean (Leonard, 1998) in the language standardized tests: expressive vocabulary (KBIT-VOC); receptive vocabulary (Peabody) and/or grammatical abilities (CEG), and/or language difficulties evidenced by oral language sample analyses using the Spanish protocol for the Evaluation of Language Delay (AREL; Pérez & Serra-Raventós, 1998). Language profiles derived from spontaneous speech transcripts provided information about children's morphosyntactic abilities in language production, and indicated a delay of at least 1 year (Bishop, 1997). Additionally, case histories completed by parents/caregivers reported concerns about oral language difficulties.

*AGE and MLU group (control groups) status confirmation*

Children in the Age and MLU groups were recruited from a public school in the metropolitan area of Barcelona. The children in the Age and MLU groups also had to score a normal non-verbal intellectual quotient (NVIQ) ≥ 70 (Nonverbal Kaufman Brief Intelligence Test [K-BIT; Kaufman & Kaufman, 2004]). For language status, participants had to demonstrate normal oral language abilities based on scores within the mean in at least two of the three standardized language tests: expressive vocabulary (KBIT-VOC), receptive vocabulary (Peabody), and grammatical abilities (CEG). Additionally, children were at grade level in school, had no history or diagnosis of language learning disability, and had never received speech and language services as reported by parents. The age-matched control group was matched in age to their DLD counterparts, while each child in the MLU group was paired with a child from the DLD group based on mean length of utterance in words (MLU-w; ±.5 words). Age, NVIQ, MLU, and language scores for each group are reported in Table 1:

*Stimuli*

Given that all three experiments share the same rationale, in this section, we describe them together in detail. The three experiments investigated the interpretation of subject pronouns in Spanish, each with a distinct methodological focus. The first two focused on overt subject pronouns with either different-or same-gender character pairs, while the third examined null pronouns. The stimuli across the three experiments followed a similar overall structure, consisting of two juxtaposed sentences paired with a visual stimulus. Experiment 1 served as a baseline and control experiment for the role of gender, in line with Arnold et al. (2000, 2007). In this experiment, the linguistic context was unambiguous, as the two human

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.



**Table 1.** Descriptives and comparison between DLD, Age and MLU groups in Age, NVIQ, MLU, and language scores

| | DLD | AGE | MLU | Comparisons between groups | | | |
|---|---|---|---|---|---|---|---|
| | *N* = 16 **M (SD)** | *N* = 16 **M (SD)** | *N* = 16 **M (SD)** | ***p* Value** | DLD vs. Age | DLD vs. MLU | AGE vs. MLU |
| **Age in months** | 93.75 (26.65) | 93.31 (28.31) | 66.56 (18.25) | <.01 | 1.00 | .01 | .01 |
| **MLU-w** | 4.72 (1.94) | 8.79 (2.44) | 5.04 (.83) | <.01 | <.01 | 1.00 | <.01 |
| **KBIT-MAT (NVIQ)** | 95.44 (14.16) | 103.37 (9.13) | 108.75 (9.54) | <.01 | .15 | <.01 | .54 |
| **PPVT-III** | 85.12 (16.72) | 104.94 (6.44) | 107.81 (10.39) | <.001 | <.001 | <.001 | 1.00 |
| **CEG** | 16.69 (17.14) | 40.56 (27.41) | 35.62 (5.86) | <.05 | <.05 | .07 | 1.00 |
| **KBIT-VOC** | 90.62 (17.23) | 108.94 (13.05) | 99.50 (13.09) | <.01 | <.01 | .27 | .22 |

*Note*: For each variable, the age-scaled scores have a mean of 100 and a SD of 15 (except CEG that are percentiles with mean 50 and age in months). MLU-w = Mean Length of Utterance by word count; NVIQ = non-verbal intellectual quotient; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; CEG = Comprehension Test of Grammatical Structures (*Test de Comprensión de Estructuras Gramaticales*); KBIT VOC = Kaufman Brief Intelligence Test—Verbal Vocabulary.

characters introduced in the first sentence differed in gender, allowing grammatical gender to function as a disambiguating cue for personal pronoun resolution. By contrast, Experiments 2 and 3 examined gender-ambiguous contexts in which both characters matched in gender, preventing gender from serving as a cue.

In the initial sentence, two human characters were introduced using definite noun phrases in subject and object positions. To prevent the most recently mentioned entity (typically the object) from influencing pronoun resolution, an inanimate element was appended through a prepositional phrase at the end of the first sentence, functioning as a buffer. The second sentence began with a subject pronoun—overt in Experiments 1 and 2 and null in Experiment 3—and included a semantically minimal verb (either *llevar* "to bring/wear" or *tener* "to have") followed by an object noun phrase. In Experiments 2 and 3, the sentence remained ambiguous until the object noun phrase of the second clause was encountered; this element, together with the accompanying image, disambiguated the intended referent of the pronoun. In contrast, Experiment 1 did not maintain this ambiguity, since gender marking on the pronoun already identified the intended referent.

To preserve referential ambiguity as much as possible prior to the onset of the critical second sentence, the verbs in the first sentence were carefully selected to minimize implicit causality biases that could favor one referent over the other. All verbs were taken from the normative ratings reported in Goikoetxea et al. (2008), and only those with near-neutral implicit causality scores were included (i.e., verbs favoring the subject between 37.5% and 62%; e.g., *asustar* "to scare," *saludar* "to greet"). By restricting the materials to verbs with minimal bias, we sought to prevent the preceding context from providing cues that might guide pronoun interpretation before participants encountered the critical elements in the following juxtaposed sentence. Because relatively few verbs met these neutrality criteria, each verb was used twice across items.

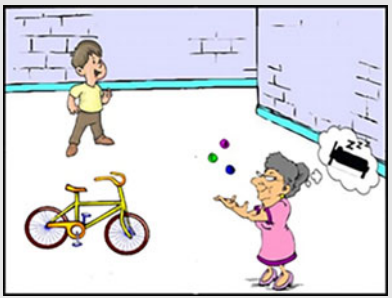
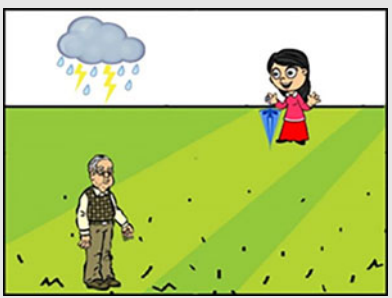
For every experiment, sixteen utterances were constructed. Experiment 1 used overt pronouns and character pairs with different grammatical genders, allowing the pronoun's referent to be unambiguously identified through gender agreement. Although the syntactic role of the antecedent (subject vs. object) varied, gender was not manipulated as an experimental variable but was instead built into the stimuli to ensure disambiguation. This experiment served as a diagnostic baseline to observe interpretation preferences in the absence of referential ambiguity. Experiments 2 and 3 introduced ambiguity by using same-sex character pairs, removing gender as a cue. In Experiment 2, overt pronouns were used, and interpretation depended solely on the syntactic role of the antecedent. Experiment 3 followed the same structure but tested null pronouns instead. In both cases, the syntactic function of the antecedent was the only variable manipulated, allowing for a focused comparison of pronoun resolution under ambiguous conditions. Table 2 outlines examples from all three experiments.

All sentences were recorded by a native Spanish-speaking female and sampled at 44.1 kHz. Recordings were edited so that the second sentence began precisely 6000 milliseconds after the start of the first, ensuring temporal consistency across items. The recordings maintained natural prosody and were perceived as unaltered by native adult listeners.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

**Table 2.** Example stimuli in each experiment

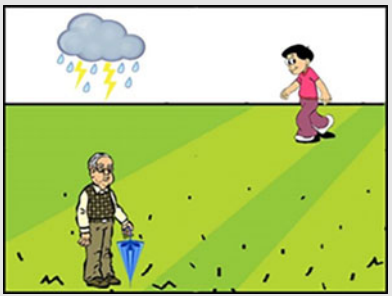
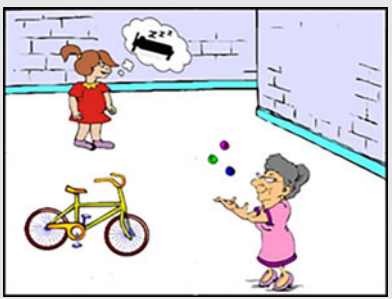
| Exp | Gender condition | Pronoun condition | Auditory item stimulus | Picture |
|---|---|---|---|---|
| 1 | Different gender | Overt-pronoun Subject antecedent | *La abuela entretiene al niño al lado de una bicicleta. Ella tiene sueño y parece que la noche tardará en llegar.*<br>"The grandmother entertains the boy next to a bicycle. She is sleepy, and it seems the night will take a while to arrive." |  |
| 1 | Different gender | Overt-pronoun Object antecedent | *El señor espera a la niña un día de tormenta. Ella lleva un paraguas y parece que lo va a necesitar*<br>"The man waits for the girl on a stormy day. She is carrying an umbrella, and it looks like he's going to need it." |  |

(*Continued*)

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.



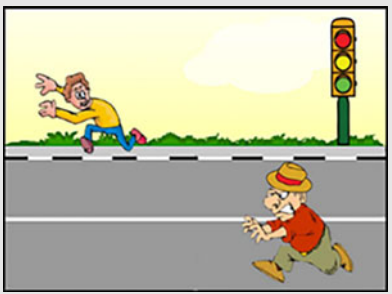
**Table 2.** (*Continued*)

| Exp | Gender condition | Pronoun condition | Auditory item stimulus | Picture |
|---|---|---|---|---|
| 2 | Same gender | Overt pronoun-Subject Antecedent | *El señor espera al niño un día de tormenta. Él lleva un paraguas y parece que lo va a necesitar*<br>"The man waits for the boy on a stormy day. He is carrying an umbrella, and it looks like he's going to need it." |  |
| 2 | Same gender | Overt-pronoun Object antecedent | *La abuela entretiene a la niña al lado de una bicicleta. Ella tiene sueño y parece que la noche tardará en llegar.*<br>"The grandmother entertains the girl next to a bicycle. She is sleepy, and it seems the night will take a while to arrive." |  |

(*Continued*)

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

**Table 2.** (*Continued*)

| Exp | Gender condition | Pronoun condition | Auditory item stimulus | Picture |
|---|---|---|---|---|
| 3 | Same gender | Null-pronoun Subject antecedent | *El señor persigue al chico cerca del semáforo. Ø Lleva un sombrero porque es carnaval y la ocasión lo merece.*<br>"The man chases the boy close to the traffic light. (He) wears a hat because it is carnival time and the occasion merits it." |  |
| 3 | Same gender | Null-Pronoun Object antecedent | *El señor persigue al chico cerca del semáforo. Ø Lleva un sombrero porque es carnaval y la ocasión lo merece.*<br>"The man chases the boy close to the traffic light. (He) wears a hat because it is carnival time and the occasion merits it." |  |

Corresponding to each sentence, a visual image was constructed and paired with each utterance. Each image consisted of a scene that was created using a clip art software and featured two characters and one distractor object (see Table 2). Sex representation was evenly distributed, with half of the items depicting female characters and the other half male. The placement of characters (target and competitor) varied between subject and object positions across items and was counterbalanced to prevent positional bias across the four screen areas, as defined by dividing the screen into quadrants (top left, top right, bottom left, and bottom right). Audio and visual elements were merged into a video file. In each video, the onset of the spoken utterance coincided with the onset of the visual stimuli. The 16 video clips created for each experiment were organized into two experimental lists (List 1 and List 2). Each list contained an equal number of trials (eight), with a balanced distribution of antecedent syntactic functions (Subject vs. Object, or First vs. Second mention). In Experiment 1, which included overt pronouns in gender-different contexts, the items differed from those used in Experiments 2 and 3 because they involved character pairs of different genders. In this case, List 1 contained items 1–8 and List 2 contained items 9–16 (see *Sentences Experiments 1, 2, and 3* in the OSF). For Experiments 2 and 3, which involved same-gender character pairs, the items were rotated across the two experiments. Specifically, in Experiment 2, List 1 contained the odd-numbered items and List 2 the even-numbered items, whereas in Experiment 3 this assignment was reversed (List 1 contained even-numbered items and List 2 odd-numbered items). Participants completed either List 1 or List 2 in both experiments, ensuring that no participant was exposed to the same lexical item or visual scene more than once (see *Sentences Experiments 1, 2, and 3* in the OSF). Additionally, each list included five filler trials—comprising both congruent ($n = 1$) and incongruent ($n = 4$) items—as well as two congruent practice items to familiarize participants with the task structure.

### *Procedure*

Children completed two testing sessions (60 minutes each) separated by. These sessions were part of a larger study, providing natural buffers between the experiments reported here. Experiment 1 was administered at the start of Session 1, followed by unrelated morphological tasks, concluding with Experiment 2. Experiment 3 was administered at the end of Session 2. Rest periods were provided throughout to minimize fatigue and carry-over effects. The time allocated to carry out the experiments of the current study was approximately 20 minutes per child. The evaluation sessions were conducted individually at the children's schools by two research assistants specifically trained for this purpose.

Participants were positioned roughly 22 inches away from a Tobii T120 eye tracker, which featured a built-in 17-inch TFT display. At the start of the experiment, a 9-point calibration was performed. Tobii Studio software automatically assessed the calibration quality, and if necessary, the experimenter could repeat the procedure to improve accuracy. The calibration process typically lasted about 20 seconds. The software was also used to display the stimuli and record eye-tracking data. Stimuli videos were created from 800 × 600 pixel images and shown on a

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

screen with a resolution of 1024 × 768 pixels. Audio associated with the stimuli was delivered through a mono signal split between two speakers placed on either side of the monitor.

During the task, participants' eye movements were tracked while they watched the visual scenes and listened to two independent sentences describing the images. Eye position data were collected at a sampling rate of 120 Hz (approximately every 8 milliseconds). Participants were asked to determine whether the story matched the image and to respond verbally with "yes" or "no." While all experimental test items were congruent, the inclusion of four incongruent filler trials per list ensured that participants could not rely on a "yes" response strategy. These filler trials served to maintain attention and verify that the children were actively processing the relationship between the audio and visual stimuli. Responses were monitored by research assistants to ensure task engagement, though they were not recorded for formal statistical analysis, as the primary measure was real-time gaze position. Before beginning the actual trials, each participant completed two practice rounds to familiarize themselves with the procedure. Participants were randomly divided into two groups, each receiving a different list of stimuli presented in randomized order. Each trial began with a fixation cross displayed for about 2000 milliseconds, which participants were instructed to focus on to ensure consistent initial gaze direction from the center of the screen.

### *Data analysis*

To characterize how gaze patterns evolved over time, we employed a non-parametric two-stage cluster-based permutation framework (Barr et al., 2014; Maris & Oostenveld, 2007), an approach used in previous studies with comparable populations and paradigms (Christou et al., 2022a, 2022b; Coloma et al., 2024; Giberga et al., 2025a, 2025b; Helo et al., 2022). This method is well-suited for autocorrelated eye-tracking time-series data because it controls the family-wise error rate without requiring independence across adjacent time bins or spatial independence between objects that are presented in the visual context at the same time. Analyses were conducted separately for the two linguistic orders—trials in which the target noun preceded the competitor and trials in which it followed.

Before the cluster analysis, raw gaze data were mapped onto two predefined areas of interest (AOIs) corresponding to the locations of the target and competitor pictures in the visual display. Fixation samples were time-locked to the onset of the critical verb phrase in the second sentence. By design, the earliest point at which referential information could influence processing differs across experiments: Experiment 1 includes a gender-marked pronoun that could trigger earlier disambiguation, Experiment 2 features an overt pronoun in a gender-neutral context, and Experiment 3 contains a null subject. To ensure comparability across experiments, all eye-tracking analyses were aligned to the verb onset in the second sentence.

Each millisecond sample was coded as a binary indicator (1 = fixation within the AOI; 0 = fixation elsewhere or track loss), and data were then aggregated into successive 50-ms time bins for each participant × item × AOI × bin cell. This means that some time bins could have an average of zero, faithfully reflecting what

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

actually happened during the trial–i.e., the absence of looks is incorporated in the proportion of fixation calculation. Thus, no trials or participants were excluded from the analyses. Descriptive statistics, including participant-level means and within-subject 95% confidence intervals adjusted following Cousineau and O'Brien (2014) and Morey (2008), are presented in the time-course figures.

As the first step of cluster permutation analysis, each 50 ms time bin was examined using a linear mixed-effects model with the lmerTest package (Kuznetsova et al., 2017) with fixation proportion as the dependent variable and Object Type as a contrast-coded fixed effect. Participant and Item were entered as random intercepts, with by-participant random slopes for Object Type to capture individual variability in the target–competitor contrast. Bins yielding a significant effect of Object Type (two-tailed $p < .05$) were flagged. Consecutive flagged bins sharing the same direction of effect and spanning a minimum of 200 ms ($\geq$4 bins) were concatenated into candidate clusters. Finally, we computed a mass statistic defined as the sum of the $t$-values across all constituent bins for each candidate cluster.

The second step involved a Monte Carlo permutation procedure. For each linguistic order condition, we generated 2000 permuted datasets by randomly shuffling the target and competitor labels within each time bin, thereby eliminating any systematic difference between objects. The full modeling and clustering pipeline was applied to each permuted dataset, yielding a null distribution of maximum cluster-mass statistics. An observed cluster was deemed significant if its mass exceeded the 97.5th percentile of the null distribution, corresponding to a two-tailed $\alpha = .05$ corrected for multiple temporal comparisons. For each significant cluster, we report the time extent, the observed cluster mass (sum of $t$-values), and the exact permutation $p$-value. Following conventions in language-mediated visual attention research, the magnitude and practical relevance of effects are characterized by three complementary indicators, the last two visible in the time-course figures: (a) the cluster mass statistic itself, which reflects both the strength and temporal persistence of the effect; (b) the vertical distance between target and competitor fixation proportion curves within significant clusters, representing the probability differential; and (c) the temporal extent of each cluster, where sustained attentional biases on the order of several hundred milliseconds are considered substantial in spoken language processing (see Barr et al., 2014).

### *Results*

Figure 1 shows the time-course of gaze behavior for the three groups—Age-controls, MLU-controls, and DLD. For every panel in the figure, solid curves depict mean fixation proportions to the target and the competitor characters, surrounded by 95% within-subject confidence-interval ribbons (shaded area); green horizontal bars identify time intervals in which cluster-based permutation tests revealed a significant preference for one character over the other. All traces are aligned to the verb onset (0 ms; *lleva* "is carrying" in Table 2), and vertical dashed lines mark the onset of the determiner (*un* "an"), the onset of the head noun (*paraguas* "umbrella"), and the offset of that noun, respectively.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

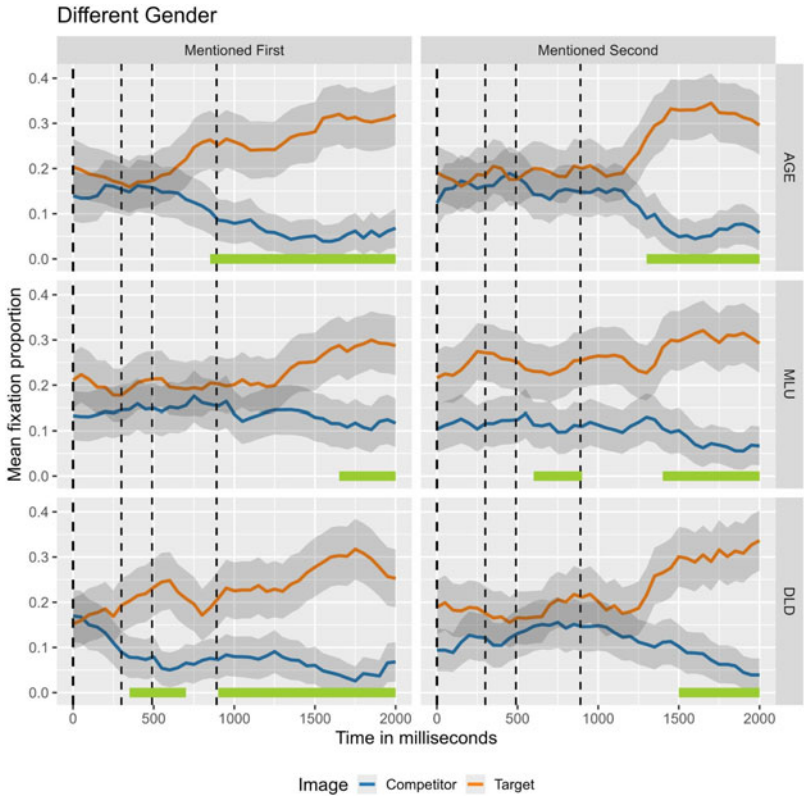



**Figure 1.** Overt pronouns in different-gender contexts.
*Note*: Proportion of looks to target (subject in first-mention condition; object in second-mention condition) and competitor (object in first-mention condition; subject in second-mention condition). On the x-axis, zero ms indicates the onset of the verb (second clause; first dashed line). The second dashed line marks the onset of the object complement; third and four dashed lines mark the onset and the offset, respectively, of the head noun object (the disambiguating word).

*AGE group.* The *mentioned-first* trials yielded a robust cluster extending from 850 to 2000 ms after verb onset (observed $t = 70.52$; $p < .001$), indicating sustained target preference. In the *mentioned-second* condition, a later cluster emerged between 1300 and 2000 ms (observed $t = 44.7$; $p < .001$), again favoring the target.

*MLU group.* No early preference was detected when the target was mentioned first, but a late cluster from 1650 to 2000 ms revealed a delayed shift toward the target (observed $t = 21.06$; $p < .001$). Conversely, *mentioned-second* trials produced an initial target preference cluster between 600 and 900 ms (observed $t = 16.16$; $p < .001$) and a second, longer cluster from 1400 to 2000 ms (observed $t = 38.69$; $p < .001$), also for the preference of the target.

*DLD group.* Two clusters characterized the *mentioned-first* condition: an early, short-lived advantage for the target from 350 to 700 ms (observed $t = 19.39$; $p < .001$), followed by a prolonged window from 900 to 2000 ms (observed $t = 63.81$; $p < .001$) where the target was preferred. In *mentioned-second* trials a

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

single late target preference cluster appeared between 1500 and 2000 ms (observed $t = 32.34$; $p < .001$).

### *Discussion*

Experiment 1 examined whether children with DLD and typically developing controls (matched on age and on MLU) could use gender cues in pronouns to resolve reference in real time. In the First-Mention condition, the DLD group displayed a surprisingly efficient use of gender cues. Their eye movements showed a sharp increase in fixations on the target referent immediately after verb offset, well before the disambiguating lexical information became available (≈0 ms vs. 500 ms). This early divergence suggests that children with DLD were able to recruit gender information in real time to anticipate the intended referent. By contrast, both control groups—age-matched and particularly the MLU-matched group—did not exhibit the same sensitivity to pronoun gender. Their fixations on the target subject only rose after the onset of the disambiguating word, with no reliable early differentiation. This finding is striking in light of the commonly reported vulnerabilities of DLD children in the domain of grammatical gender and gender agreement (Bedore & Leonard, 2001; Leonard, 2014). A plausible explanation is that the cues at stake in this task were not instances of arbitrary grammatical gender but rather of semantic gender, grounded in the sex of animate and human referents. As Corbett (1991) argues, such cases differ in nature from purely formal gender categories because they map directly onto conceptual knowledge. The present results suggest that DLD children may actually benefit from the salience of sex-based distinctions, which provide a reliable semantic anchor that supports rapid reference resolution, while control children, and especially younger MLU controls, did not exploit these cues as effectively in real time.

The Second-Mention condition, which posed a greater processing challenge, further differentiated the groups. The DLD group showed some early sensitivity to gender cues, as indicated by a modest divergence in fixations shortly after verb offset. However, they were unable to fully resolve the referent until after the disambiguating information was available. The MLU group displayed an early divergence in fixations toward the target referent just before the onset of the disambiguating lexical cue. In contrast, the age-matched control group showed no evidence of using gender cues—either early or later—with resolution occurring only well after the offset of the disambiguating word. The DLD group exhibited an intermediate pattern between the two control groups: although they showed a weak early tendency toward the target referent, successful resolution occurred only after disambiguation, slightly later than in the MLU group and later still than in the age-matched group. These findings suggest that some children with DLD can process gender-related information but do so less efficiently and with slower real-time integration.

Taken together, the findings from Experiment 1 indicate that while typically developing children—particularly those matched on age—cannot effectively use gender cues in overt pronouns to guide reference resolution, children with DLD show some ability to exploit such cues, clearly grounded in sex contrasts. In unambiguous contexts where gender clearly signals the intended referent, children

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

with DLD seem able to integrate this semantic information during online processing. This efficiency likely contributes to mitigating some of their difficulties in discourse comprehension and anaphora resolution. More broadly, this pattern resonates with psycholinguistic evidence that listeners are generally faster and more accurate when processing cues that are conceptually transparent and semantically predictable (e.g., animacy hierarchies, thematic role plausibility).

The ability of children with DLD to take advantage of semantic gender thus highlights a relative processing strength: when morphosyntactic information is reinforced by conceptual salience, their performance can even surpass that of typically developing peers matched for age or MLU. Such results underscore the importance of distinguishing between formal, arbitrary gender systems and semantically grounded gender in understanding both the challenges and compensatory strategies of children with DLD.

## Experiment 2

This experiment focused on the resolution of overt pronouns in ambiguous same-gender contexts.

### Method

*Participants*

The same participants who took part in Experiment 1 took part in Experiment 2.

*Stimuli*

Details can be found in Experiment 1.

*Procedure*

The same procedure as in Experiment 1 was used.

*Data analysis*

The same procedure as in Experiment 1 was used.

### Results

Figure 2 presents the gaze-contingent time courses for the three child groups. As in Experiment 1, solid lines represent mean fixation proportions to the target and the competitor, with 95% within-subject confidence-interval ribbons; green horizontal bars indicate periods in which cluster-based permutation tests detected a significant object preference. The data were time-locked to the verb onset (ms; *lleva* "is carrying" in Table 2), and vertical dashed lines successively mark the determiner onset (*un* "an"), the head-noun onset (*paraguas* "umbrella"), and the noun offset.

*AGE group.* When the target was mentioned first, a single late cluster emerged from 1500 to 2000 ms after verb onset, signaling a sustained bias toward the target (observed $t = 30.12$; $p < .001$). The mentioned-second condition yielded a comparable but slightly earlier cluster spanning 1450 to 2000 ms, again favoring the target (observed $t = 34.03$; $p < .001$).

*MLU group.* Only the mentioned-first condition elicited a significant cluster (1050–2000 ms, observed $t = 50.59$; $p < .001$), revealing an extended late shift toward the target. No reliable clusters were detected when the target was mentioned

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

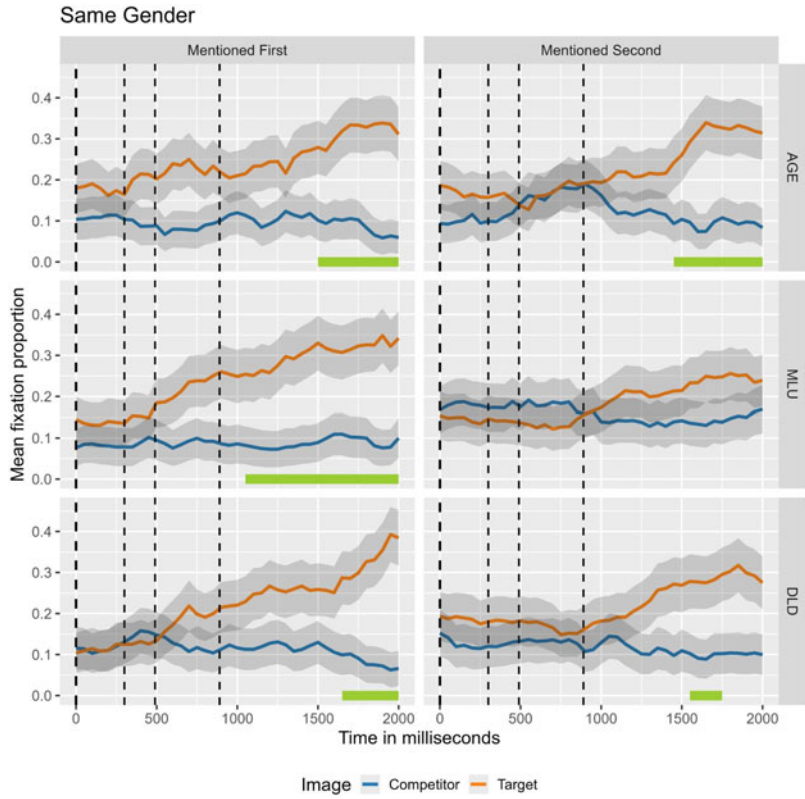



**Figure 2.** Overt pronouns in same-gender contexts.
*Note:* Proportion of looks to target (subject in first-mention condition; object in second-mention condition) and competitor (object in first-mention condition; subject in second-mention condition). Zero is the onset of the pronoun. The second dashed line marks the onset of the object complement; third and four dashed lines mark the onset and the offset, respectively, of the head noun object (the disambiguating word).

second, indicating the absence of systematic preference under that ordering for this group.

*DLD group.* In the mentioned-first trials, the DLD children displayed a delayed yet prolonged cluster between 1650 and 2000 ms (observed $t = 23.31$; $p < .001$) with a preference for the target. By contrast, the mentioned-second trials produced a shorter cluster confined to 1550–1750 ms, suggesting a more transient discrimination when the target followed the competitor (observed $t = 12.97$; $p < .001$).

### *Discussion*

Experiment 2 explored how children interpret third-person subject pronouns when both antecedents share the same grammatical gender. In this context, children could not rely on semantic cues like semantic gender and instead had to draw on discourse-level cues such as order of mention and grammatical role. This allowed us

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

to investigate the extent to which children with DLD, children with typical development, and children matched by MLU-w can use these higher-level cues to resolve pronominal reference.

In the First-Mention condition, all three groups eventually showed a bias toward the first-mentioned referent, as reflected in a cluster indicating a preference for the target over the competitor. However, this preference emerged at different time points across groups. The earliest cluster appeared in the MLU group (around 1050 ms), followed by the Age group (around 1500 ms), and finally the DLD group (around 1650 ms). Thus, although participants showed sensitivity to the discourse prominence of the first-mentioned referent, this preference did not arise immediately after pronoun onset. Importantly, none of the groups fully resolved the referent until disambiguating information became available. This is reflected in the fact that differences between looks to the target and the competitor only reached statistical significance after the disambiguating lexical cue was presented, as indicated by the green horizontal bars marking significant preferences for one character over the other. These results suggest that while children are sensitive to discourse prominence cues such as first mention, the successful resolution of pronoun reference continues to rely on explicit linguistic information. The overall tendency to interpret ambiguous pronouns as referring to first-mentioned entities is consistent with previous findings in typically developing children as well as Spanish adults, which show that first-mentioned or subject antecedents are more accessible in memory and therefore more likely to be selected as referents. Interestingly, the DLD group also showed this early subject bias, favoring the first-mentioned referent almost immediately after pronoun onset. However, they were the slowest group to clearly resolve the pronoun reference, doing so well after the disambiguating information—later than both the Age and MLU groups.

In the Second-Mention condition, children faced a greater interpretive challenge, as the correct referent was not the first-mentioned entity. In this context, none of the groups showed a clear preference for either referent until well after the offset of the disambiguating word. Children with DLD displayed a pattern similar to that of the Age group, showing a very late preference toward the target referent. In contrast, the MLU group did not show a reliable preference for either the target or the competitor at any point during the resolution process, maintaining a more underspecified interpretation until the end of the sentence, which suggests greater difficulty with pronoun resolution in this condition.

The results suggest that children in all groups exhibited a bias toward the first-mentioned antecedent when interpreting pronouns, particularly in cases where both potential referents shared the same grammatical gender, and the pronoun was explicit, yet full resolution occurred only once disambiguating information was provided.

## Experiment 3

In this experiment, we shifted our focus from overt to null pronouns, investigating their resolution in ambiguous same-gender contexts.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

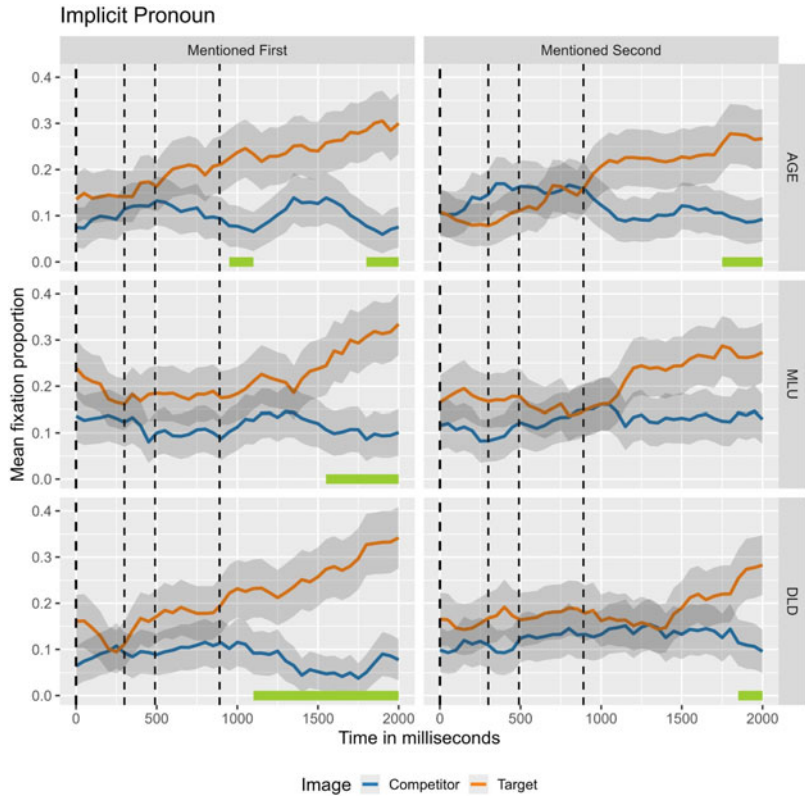



**Figure 3.** Null pronouns in same-gender contexts.
*Note:* Proportion of looks to target (subject in first-mention condition; object in second-mention condition) and competitor (object in first-mention condition; subject in second-mention condition). On the x-axis, 0 ms indicates the onset of the verb (the beginning of the second clause; marked by the first dashed line). The second dashed line marks the onset of the object complement; third and four dashed lines mark the onset and the offset, respectively, of the head noun object (the disambiguating word).

## Method

*Participants*

The same participants that in Experiment 1 took part in Experiment 3.

*Stimuli*

Details can be found in Experiments 1 and 2.

*Procedure*

The same procedure as in Experiments 1 and 2 was used.

*Data analysis*

The same procedure as in Experiments 1 and 2 was used.

## Results

Figure 3 depicts the fixation time courses for Age-controls, MLU-controls, and DLD children under the two noun-order conditions. As in the previous experiments,

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

solid curves show mean fixation proportions to the target and competitor, shaded areas around them represent 95% within-subject confidence interval; colored horizontal bars mark the extent of the cluster that reached a significant object preference. As in previous plots, data are aligned to verb onset (0 ms; *lleva* "is wearing" in Table 2), with vertical dashed lines denoting, in order, determiner onset (*un* "a"), head-noun onset (*sombrero* "hat"), and noun offset.

*AGE group.* When the target was mentioned first, two distinct clusters emerged: an early window from 950 to 1100 ms (observed $t = 9.34$; $p < .001$) and a late, longer window from 1800 to 2000 ms after verb onset (observed $t = 12.91$; $p < .001$), demonstrating a biphasic shift toward the target. In the mentioned-second condition, a single late cluster spanned 1750–2000 ms (observed $t = 14.85$; $p < .001$), mirroring the sustained late bias to the target object seen in Experiments 1 and 2.

*MLU group.* A single late cluster appeared when the target preceded the competitor, beginning at 1550 ms and persisting to 2000 ms (observed $t = 24.82$; $p < .001$), revealing a delayed but sustained shift toward the target. No reliable clusters were observed when the target was mentioned second, again reflecting the group's limited ability to establish referential certainty under that ordering.

*DLD group.* Mentioned-first trials produced one prolonged cluster extending from 1100 to 2000 ms (observed $t = 54.52$; $p < .001$), signaling a delayed yet steady preference for the target. In the mentioned-second condition, a shorter late target preference cluster was detected between 1850 and 2000 ms (observed $t = 9.73$; $p < .001$), indicating that reference resolution was achieved only near the end of the time window.

### Discussion

Experiment 3 examined how children interpret null third-person subject pronouns in Spanish when both antecedents shared the same grammatical gender. Since Spanish allows subject pronouns to be omitted, listeners must rely more heavily on discourse-level cues, such as order of mention and grammatical role, to resolve reference. This experiment thus tested children's ability to interpret null pronouns using structural information alone, without the support of overt semantic or morphological cues.

In the First-Mention condition, all three groups showed a tendency to favor the first-mentioned referent, reflecting a subject bias in the interpretation of the null pronoun. However, this preference only emerges after the disambiguating word region has been reached, suggesting delayed bias in the time course of processing. Overall, these results indicate that children across groups were sensitive to discourse-level cues—such as subjecthood and linear order—when interpreting the null pronoun. However, as in Experiment 2, they were unable to fully resolve the referent until after the disambiguating information was available; nonetheless, a preference for the subject was maintained during the ambiguous region, as indicated by the proportion of fixations (with the orange line consistently remaining above the blue line). The DLD group's performance in this condition was particularly noteworthy: their early and sustained preference for the first-mentioned referent mirrored that of the typically developing age group. In contrast, the MLU group

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

exhibited a delayed and more uncertain response. Their fixations were relatively flat until approximately 1250–1300 milliseconds after verb offset, indicating a delayed commitment to a referent. This finding supports the pattern attested in Experiment 2 with overt pronouns and suggests that younger children matched on expressive language level are still developing the ability to integrate structural cues efficiently during online processing, even when semantic information is not required.

The Second-Mention condition presented a greater challenge, as it required children to override the default subject bias and assign the pronoun to the second-mentioned entity. In this condition, the age group initially showed a subject preference, but began to shift toward the correct referent just before the offset of the disambiguating word—around 1000 milliseconds. This suggests that even typically developing children needed time and additional contextual information to override the strong, inherent subject bias of null pronouns when it was not consistent with the intended referent. The MLU group showed an interesting pattern: they exhibited an early preference for the second-mentioned referent, possibly reflecting recency effects as a consequence of not yet having developed a subject antecedent bias with null pronouns, as noted above. However, this early preference did not translate into consistent referent identification until just after the disambiguating information was fully processed. The DLD group showed the most delayed response in this condition. They did not display a clear preference for the appropriate referent until well after the disambiguating word had been fully processed—around 1250 to 1300 milliseconds—suggesting difficulty in suppressing the subject bias when it conflicted with the intended interpretation. This pattern further underscores a core limitation in the DLD group's processing: while they can successfully rely on structural defaults when those align with the correct selection, they struggle to revise or override initial interpretations in more demanding conditions that require flexible use of contextual information.

In sum, the findings from Experiment 3 reinforce and extend those of Experiment 2. Both experiments reveal a subject or first-mentioned bias across all groups, albeit less evident in the MLU-w group, indicating that the first-mentioned discourse strategy is still developing. However, while the DLD group performed similarly to the Age group when the correct referent matched their default strategy (i.e., in the First-Mention condition), they exhibited marked difficulties when this strategy led to an inaccurate interpretation (i.e., in the Second-Mention condition). These results highlight a key characteristic of sentence processing in children with DLD: a reliance on default cues based on discourse-syntactic structure paired with a limited ability to revise initial interpretations when more complex discourse integration is required.

## General discussion

This study investigated how Spanish-speaking children with DLD interpret third-person subject pronouns during real-time sentence processing. By employing the Visual World Paradigm across three experiments, we aimed to determine how these children weigh semantic and structural cues. Below, we revisit our initial hypotheses and discuss the theoretical implications of the observed bias patterns.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

### Semantic gender cues (Experiment 1)

Our first hypothesis predicted that children with DLD would successfully exploit sex-based gender cues to resolve pronoun referents in unambiguous contexts. This hypothesis was confirmed. The results demonstrate that children with DLD can effectively use grammatical gender when antecedents differ in biological sex. In fact, this effect was more pronounced in the DLD group than in TD controls, suggesting an over-reliance on salient conceptual "anchors" to compensate for broader processing difficulties. However, the fragility of this effect in the Second-Mention condition indicates that while the information is available, the processing remains resource-dependent.

### Structural cues in overt pronouns (Experiment 2)

Based on prior evidence of reduced accuracy in DLD (Rakhlin et al., 2015), we predicted that children with DLD would not exhibit a consistent first-mention or subject-preference bias. This hypothesis was disconfirmed. Contrary to our expectations, children with DLD displayed a clear, early bias toward the first-mentioned subject, mirroring TD patterns. This suggests the "subject-bias" heuristic is fundamentally intact, though the DLD group lacked the flexibility to revise this default when faced with conflicting late-arriving cues.

### Structural cues in null pronouns (Experiment 3)

We posited that children with DLD would show inconsistent use of structural cues for null pronouns. This hypothesis was partially confirmed. While the DLD group showed a robust early subject bias in congruent (First-Mention) contexts, they displayed significantly delayed and inconsistent resolution in incongruent (Second-Mention) contexts. This confirms that the difficulty lies not in the presence of the default bias, but in the sophisticated integration required to manage non-canonical discourse.

### The overt vs. null contrast: subject bias in real-time

A particularly striking finding in this study—and one that warrants further clarification—is the observation of a subject/first-mention bias for both overt (Exp 2) and null (Exp 3) pronouns. In null-subject languages like Spanish, the Position of Antecedent Strategy (PAS) (e.g., Sorace et al., 2009) typically predicts a division of labor: null pronouns prefer subject antecedents, while overt pronouns (*él, ella*) prefer non-subjects (e.g., objects). However, our real-time data suggests that at the earliest stages of processing, both forms initially trigger a search for the most prominent discourse entity (the subject).

This apparent contrast with the PAS can be explained by the difference between initial processing heuristics and final interpretive conventions. Our results suggest that:

1. General Prominence: At the moment of pronoun encounter, children (both TD and DLD) apply a "prominence heuristic" (Song & Fisher, 2005) that

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

favors the subject/first-mention, regardless of the pronoun's phonetic realization.
2. Delayed Differentiation: The language-specific pragmatic convention (where overt pronouns signal a topic shift to the non-subject) may emerge later in the processing window or require more cognitive resources to implement than the universal subject-preference.

For children with DLD, this results in a "double-edged sword." Their intact subject-bias allows them to process canonical null-subject sentences efficiently, but their rigid adherence to this bias makes the non-subject preference typically associated with overt pronouns much harder to access in real time.

### *Synthesis and theoretical implications*

Taken together, these findings reveal a consistent pattern: children with DLD are not "cue-blind." Instead, they demonstrate a rigid reliance on default heuristics (gender congruence and subjecthood). While such heuristics are foundational to the probabilistic interpretive system used by TD children (Blything et al., 2022), the children with DLD in our study showed a marked inability to inhibit or revise these defaults.

This rigid processing style differs from the findings of Van der Lely and Stollwerck (1997) and Rakhlin et al. (2015). We attribute this divergence to our use of real-time eye-tracking. While offline tasks often highlight a total failure in grammatical principles, our online data shows that the initial biases are present, but the integration and revision processes are what ultimately fail. This suggests that the core deficit in DLD pronoun resolution is a limitation in the dynamic, real-time integration of multiple linguistic and cognitive resources.

## Limitations and future directions

This study explored how children with DLD interpret pronouns in real time using eye-tracking within the Visual World Paradigm. Nonetheless, there is room for future improvement. Although efforts were made to recruit a large number of participants, the sample included 48 children, with 16 diagnosed with DLD, 16 typically developing age-matched children, and 16 younger children matched by MLU-w. The sample size in the present study aligns with those reported in earlier research on pronoun resolution in children with DLD (Abbott et al., 2024; Rakhlin et al., 2015; Van der Lely & Stollwerck, 1997). However, our study included participants ranging from 4;00 to 12;03 years, representing a diverse age range. Although previous studies also included participants with a wide range of ages across all groups—4;00 to 12;03 years (Van der Lely & Stollwerck, 1997), 5;01 to 10;11 years (Rakhlin et al., 2015), and 6;02 to 11;08 years (Abbott et al., 2024)—using narrower age ranges and larger sample sizes could yield more precise information on the development of pronoun reference assignment.

Another limitation is the absence of explicit offline measures assessing pronoun antecedent identification. As a result, we have visual data reflecting the process of pronoun resolution but no explicit information about participants' final decisions.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

An explicit task like the one used in the second experiment of Abbott et al. (2024) would reveal children's final pronoun reference assignment. Such data could provide insight into the ultimate interpretation of the pronoun antecedent, particularly in cases where the eye-tracking results do not offer a clear attribution. Future research should combine online and offline measures to examine the moment-by-moment processing of pronouns while also determining the ultimate assignment of pronoun reference.

**Replication package.** The materials, data, and scripts used in this study are available on the Open Science Framework (OSF) at the following link: https://osf.io/gh68q/.

**Acknowledgments.** This work was supported by the Spanish State Research Agency (grant numbers EDU2013-44678-P and PID2020-114690RB-I00, PI: Llorenç Andreu; FFI2016-75082-P and FFI2012-35058, PI: Aurora Bel), and by the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, Generalitat de Catalunya) (grant numbers 2021 SGR 00472; 2017 SGR 1028; 2021 SGR 01102; 2017 SGR 387). Funding from ANID/Support 2024AFB240004 (to Ernesto Guerra) is gratefully acknowledged. Funding from ANID CIAE CIA250005 is gratefully acknowledged.

## References

**Abbott, N., Nip, I., & Love, T.** (2024). Rate of speech affects the comprehension of pronouns in children with developmental language disorder. *Frontiers in Language Sciences*, **3**, 1394742. https://doi.org/10.3389/flang.2024.1394742.

**Abel, A. D., Rice, M. L., & Bontempo, D. E.** (2015). Effects of verb familiarity on finiteness marking in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **58**(2), 360–372. https://doi.org/10.1044/2015_JSLHR-L-14-0003.

**Arnold, J. E., Brown-Schmidt, S., & Trueswell, J.** (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, **22**(4), 527–565. https://doi.org/10.1080/01690960600845950.

**Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C.** (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, **76**(1), B13–B26. https://doi.org/10.1016/S0010-0277(00)00073-1.

**Bader, M., & Portele, Y.** (2019). The interpretation of German personal pronouns and d-pronouns. *Zeitschrift für Sprachwissenschaft*, **38**(2), 155–190. https://doi.org/10.1515/zfs-2019-2002.

**Barr, D. J., Jackson, L., & Phillips, I.** (2014). Using a voice to put a name to a face: The psycholinguistics of proper name comprehension. *Journal of Experimental Psychology: General*, **143**(1), 404. https://doi.org/10.1037/a0031813.

**Bedore, L. M., & Leonard, L. B.** (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **44**(4), 905–992. https://doi.org/10.1044/1092-4388(2001/072).

**Bel, A., & Albert, M.** (2016). The Development of Referential Choice in Spanish Narratives Among School-Age Children and Adolescents. In J. Perera, M. Aparici, E. Rosado, & N. Salas, (Eds.), *Written and spoken language development across the lifespan. Essays in honour of Liliana Tolchinsky* (pp. 251–269). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-21136-7_15.

**Bel, A., Guerra, E., Ahufinger, N., Andreu, L., & Sanz-Torrent, M.** (2026). Resolution of Null and overt pronouns in Catalan: An eye tracking study. *Journal of Psycholinguistic Research*, **55**, 7. https://doi.org/10.1007/s10936-025-10189-7.

**Bel, A., Perera, J., & Salas, N.** (2010). Anaphoric devices in written and spoken narrative discourse: Data from Catalan. *Written Language & Literacy*, **13**(2), 236–259. https://doi.org/10.1075/wll.13.2.03bel.

**Bishop, D.** (1997). *Uncommon Understanding*. Psychology Press.

**Bishop, D. V.** (2003). *Test for Receptive Grammar, TROG-2*. Pearson.

**Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T.**, & **Catalise-2 Consortium**. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

language development: Terminology. *Journal of Child Psychology and Psychiatry*, **58**(10), 1068–1080. https://doi.org/10.1111/jcpp.12721.

**Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Catalise Consortium**. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS One*, **11**(7), e0158753. https://doi.org/10.1371/journal.pone.0158753.

**Blything, L. P., Azpiroz, M. I., Allen, S., Hert, R., & Järvikivi, J.** (2022). The influence of prominence cues in 7-to 10-year-olds' pronoun resolution: Disentangling order of mention, grammatical role, and semantic role. *Journal of Child Language*, **49**(5), 930–958. https://doi.org/10.1017/S0305000921000349.

**Bortolini, U., Arfé, B., Caselli, C. M., Degasperi, L., Deevy, P., & Leonard, L. B.** (2006). Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language & Communication Disorders*, **41**(6), 695–712. https://doi.org/10.1080/13682820600570831.

**Calder, S. D., Brennan-Jones, C. G., Robinson, M., Whitehouse, A., & Hill, E.** (2022). The prevalence of and potential risk factors for developmental language disorder at 10 years in the Raine Study. *Journal of Paediatrics and Child Health*, **58**(11), 2044–2050. https://doi.org/10.1111/jpc.16149.

**Carreiras, M., Garnham, A., & Oakhill, J.** (1993). The use of superficial and meaning-based representations in interpreting pronouns: Evidence from Spanish. *European Journal of Cognitive Psychology*, **5**(1), 93–116. https://doi.org/10.1080/09541449308406516.

**Christou, S., Andreu, L., Coloma, C. J., Guerra, E., Araya, C., Rodriguez-Ferreiro, J., & Sanz-Torrent, M.** (2022a). Insights from real-time comprehension of Spanish verbal tense in children with developmental language disorder: An eye-tracking study. *Applied Psycholinguistics*, **43**(3), 641–662. https://doi.org/10.1017/S0142716422000042.

**Christou, S., Coloma, C. J., Andreu, L., Guerra, E., Araya, C., Rodriguez-Ferreiro, J., & Sanz-Torrent, M.** (2022b). Online comprehension of verbal number morphology in children with developmental language disorder: An eye-tracking study. *Journal of Speech, Language, and Hearing Research*, **65**(11), 4181–4204. https://doi.org/10.1044/2022_JSLHR-21-0059.

**Cipriani, P., Bottari, P., Chilosi, A., & Pfanner, L.** (1998). A longitudinal perspective on the study of specific language impairment: The long term follow-up of an Italian child. *International Journal of Language & Communication Disorders*, **33**(3), 245–280. https://doi.org/10.1080/136828298247749.

**Coloma, C. J., Araya, C., Quezada, C., Pavez, M. M., & Maggiolo, M.** (2016). Grammaticality and complexity of sentences in monolingual Spanish-speaking children with specific language impairment. *Clinical Linguistics & Phonetics*, **30**(9), 649–662. https://doi.org/10.3109/02699206.2016.1163420.

**Coloma, C. J., Guerra, E., De Barbieri, Z., & Helo, A.** (2024). Article comprehension in monolingual Spanish-speaking children with developmental language disorder: A longitudinal eye tracking study. *International Journal of Speech-Language Pathology*, **26**(1), 105–117. https://doi.org/10.1080/17549507.2023.2167235.

**Colonna, S., Schimke, S., & Hemforth, B.** (2015). Different effects of focus in intra-and inter-sentential pronoun resolution in German. *Language, Cognition and Neuroscience*, **30**(10), 1306–1325. https://doi.org/10.1080/23273798.2015.1066510.

**Corbett, G. G.** (1991). *Gender*. Cambridge University Press.

**Cousineau, D., & O'Brien, F.** (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, **46**(4), 1149–1151. https://doi.org/10.3758/s13428-013-0441-z.

**Crawley, R. A., Stevenson, R. J., & Kleinman, D.** (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, **19**(4), 245–264. https://doi.org/10.1007/BF01077259.

**de la Fuente, I.** (2015). Putting pronoun resolution in context: The role of syntax, semantics, and pragmatics in pronoun interpretation [Doctoral dissertation, Université Paris Diderot]. HAL Theses. https://tel.archives-ouvertes.fr/tel-01535977.

**Dickinson, K. V., Ortiz-Ramírez, P. A., Arrieta-Zamudio, A., Grinstead, J., & Flores-Ávalos, B.** (2023). Overt subject pronoun use in switch-reference contexts in Child Spanish developmental language disorder: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research*, **66**(2), 605–619. https://doi.org/10.1044/2022_JSLHR-21-00483.

**Dunn, Ll. M; Dunn, L. M., & Arribas. D.** (2006). *Test de vocabulario en imágenes (PEABODY: PPVT-III)*. TEA Ediciones.

**Filiaci, F., & Baldo, M.** (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, **119**(3), 460–477. https://doi.org/10.1016/j.lingua.2008.09.008.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

**Giberga, A., Guerra, E., Ahufinger, N., Igualada, A., Aguilera, M., & Esteve-Gibert, N.** (2025a). How children with and without developmental language disorder use prosody and gestures to process phrasal ambiguities. *Languages*, **10**(4), 61. https://doi.org/10.3390/languages10040061.

**Giberga, A., Guerra, E., Ahufinger, N., Igualada, A., Aguilera, M., & Esteve-Gibert, N.** (2025b). Prosody and gestures help pragmatic processing in children with developmental language disorder. *Journal of Communication Disorders*, **115**, 106525. https://doi.org/10.1016/j.jcomdis.2025.106525

**Goikoetxea, E., Pascual, G., & Acha, J.** (2008). Normative study of the implicit causality of 100 interpersonal verbs in Spanish. *Behavior Research Methods*, **40**, 760–772. https://doi.org/10.3758/BRM.40.3.760.

**Grinstead, J., Lintz, P., Pratt, A., Vega-Mendoza, M., De la Mora, J., Cantú-Sánchez, M., & Flores-Avalos, B.** (2018). Null subject occurrence in monolingual Spanish SLI: A discriminant function analysis. *Languages*, **3**(2), 17. https://doi.org/10.3390/languages3020017

**Grüter, T.** (2006). Object clitics and null objects in the acquisition of French (Doctoral dissertation, McGill University, Department of Linguistics).

**Helo, A., Guerra, E., Coloma, C. J., Reyes, M. A., & Rämä, P.** (2022). Objects shape activation during spoken word recognition in preschoolers with typical and atypical language development: An eye-tracking study. *Language Learning and Development*, **18**(3), 324–351.

**Hendriks, P., Koster, C., & Hoeks, J. C.** (2014). Referential choice across the lifespan: Why children and elderly adults produce ambiguous pronouns. *Language, Cognition and Neuroscience*, **29**(4), 391–407. https://doi.org/10.1080/01690965.2013.766356.

**Hert, R., Järvikivi, J., & Arnhold, A.** (2024). The importance of linguistic factors: He likes subject referents. *Cognitive Science*, **48**(4), e13436. https://doi.org/10.1111/cogs.13436.

**Hickmann, M.** (2004). Coherence, Cohesion, and Context: Some Comparative Perspectives in Narrative Development. In S. Strömqvist, & L. Verhoeven (Eds.), *Relating events in narrative: Vol. 2. Typological and contextual perspectives* (pp. 281–306). Psychology Press.

**Jakubowicz, C., Nash, L., Rigaut, C., & Gerard, C. L.** (1998). Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition*, **7**(2–4), 113–160. https://doi.org/110.1207/s15327817la0702-4_3.

**Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B.** (2014). Information structure cues for 4-year-olds and adults: Tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience*, **29**(7), 877–892. https://doi.org/10.1080/01690965.2013.804941.

**Järvikivi, J., Van Gompel, R. P., Hyönä, J., & Bertram, R.** (2005). Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. *Psychological Science*, **16**(4), 260–264. https://doi.org/10.1111/j.0956-7976.2005.01525.x.

**Kaiser, E., & Trueswell, J.** (2003). Putting things in context: Sentence processing in languages with flexible word order. U. Penn Working Papers in Linguistics, *9.1*, 105–118.

**Káldi, T., & Babarczy, A.** (2021). Linguistic focus guides attention during the encoding and refreshing of working memory content. *Journal of Memory and Language*, **116**, 104187. https://doi.org/10.1016/j.jml.2020.104187.

**Kaufman, A. S., & Kaufman, N. L.** (2004). K-BIT: Test breve de inteligencia Kaufman.

**Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H.** (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, **82**, 1–26. https://doi.org/10.18637/jss.v082.i13.

**Leonard, L. B.** (1998). *Language, speech, and communication. Children with specific language impairment*. The MIT Press.

**Leonard, L. B.** (2014). *Children with specific language impairment*. MIT Press.

**Loeb, D. F., & Leonard, L. B.** (1991). Subject case marking and verb morphology in normally developing and specifically language-impaired children. *Journal of Speech, Language, and Hearing Research*, **34**(2), 340–346. https://doi.org/10.1044/jshr.3402.340.

**MacWhinney, B., & Bates, E.** (1989). *The crosslinguistic study of sentence processing*. Cambridge University Press.

**Maris, E., & Oostenveld, R.** (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, **164**(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

**McDonald, J. L., & MacWhinney, B.** (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, **34**(4), 543–566. https://doi.org/10.1006/jmla.1995.1025.

**Mendoza, E., Carballo G., Muñoz, J., & Fresneda, M. D.** (2006). *Test de Comprensión de Estructuras Gramaticales*. TEA Ediciones.

**Moore, M. E.** (1995). Error analysis of pronouns by normal and language-impaired children. *Journal of Communication Disorders*, **28**(1), 57–72. https://doi.org/10.1016/0021-9924(94)00007-M.

**Moore, M. E.** (2001). Third person pronoun errors by children with and without language impairment. *Journal of Communication Disorders*, **34**(3), 207–228. https://doi.org/10.1016/S0021-9924(00)00050-2.

**Morey, R. D.** (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, **4**(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061.

**Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A.** (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, **57**(11), 1247–1257. https://doi.org/10.1111/jcpp.12573.

**Paradis, J., Crago, M., & Genesee, F.** (2006). Domain-specific versus domain-general theories of the deficit in SLI: Object pronoun acquisition by French–English bilingual children. *Language Acquisition*, **14**, 33–62. https://doi.org/10.1207/s15327817la1301_3

**Patterson, C., Esaulova, Y., & Felser, C.** (2017). The impact of focus on pronoun resolution in native and non-native sentence comprehension. *Second Language Research*, **33**(4), 403–429. https://doi.org/10.1177/026765831769778.

**Pérez, E., & Serra-Raventós, M.** (1998). *Análisis del retraso del lenguaje: AREL*. Ariel.

**Rakhlin, N., Kornilov, S. A., Reich, J., & Grigorenko, E. L.** (2015). Interpretation of Anaphoric dependencies in Russian-speaking children with and without developmental language disorder. *Language Acquisition*, **22**(4), 355–383. https://doi.org/10.1080/10489223.2015.1028629.

**Restrepo, M. A., & Gutierrez-Clellen, V. F.** (2001). Article use in Spanish-speaking children with specific language impairment. *Journal of Child Language*, **28**(2), 433–452. https://doi.org/10.1017/S0305000901004706.

**Sanz-Torrent, M., Serrat, E., Andreu, L., & Serra, M.** (2008). Verb morphology in Catalan and Spanish in children with specific language impairment: A developmental study. *Clinical Linguistics & Phonetics*, **22**(6), 459–474. https://doi.org/10.1080/02699200801892959.

**Schul, R., Stiles, J., Wulfeck, B., & Townsend, J.** (2004). How "generalized" is the "slowed processing" in SLI?: The case of visuospatial attentional orienting. *Neuropsychologia*, **42**, 661–671. https://doi.org/10.1016/j.neuropsychologia.2003.10.010.

**Serratrice, L.** (2007). Referential cohesion in the narratives of bilingual English-Italian children and monolingual peers. *Journal of Pragmatics*, **39**(6), 1058–1087. https://doi.org/10.1016/j.pragma.2006.10.001.

**Shin, N. L., & Cairns, H. S.** (2012). The development of NP selection in school-age children: Reference and Spanish subject pronouns. *Language Acquisition*, **19**(1), 3–38. https://doi.org/10.1080/10489223.2012.633846.

**Song, H. J., & Fisher, C.** (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, **52**(1), 29–57. https://doi.org/10.1016/j.jml.2004.06.012.

**Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M.** (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, **119**(3), 460–477.

**Stavrakaki, S.** (1999). The Case of a Greek SLI Child: Evaluation of Person and Number Agreement, Case Assignment to Overt Subject Pronouns and Tense Marking. In A. Kakouriotis, & V. BollaMavrides (Eds.), *Working papers in theoretical and applied linguistics* (Vol. **6**, pp. 165–187). Thessaloniki.

**Stavrakaki, S., & Van der Lely, H.** (2010). Production and comprehension of pronouns by Greek children with specific language impairment. *British Journal of Developmental Psychology*, **28**(1), 189–216. https://doi.org/10.1348/026151010X485269.

**Stevenson, R. J., Crawley, R. A., & Kleinman, D.** (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, **9**(4), 519–548. https://doi.org/10.1080/01690969408402130.

**Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C.** (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, **268**(5217), 1632–1634. https://10.1126/science.7777863.

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.

**Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M.** (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, **40**(6), 1245–1260. https://doi.org/10.1044/jslhr.4006.1245.

**Trueswell, J. C.** (2008). Using Eye Movements as a Developmental Measure within Psycholinguistics. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Language processing in children* (pp. 73–96). John Benjamins.

**Van der Lely, H. K., & Stollwerck, L.** (1997). Binding theory and grammatical specific language impairment in children. *Cognition*, **62**(3), 245–290. https://doi.org/10.1016/S0010-0277(96)00783-4.

**Van der Lely, H. K., & Ullman, M. T.** (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, **16**(2–3), 177–217. https://doi.org/10.1080/01690960042000076.

**Cite this article:** Andreu, L., Bel, A., Guerra, E., Ahufinger, N., & Sanz-Torrent, M. (2026). Online pronoun resolution in children with developmental language disorder (DLD): A visual world eye-tracking study. *Applied Psycholinguistics*. https://doi.org/10.1017/S014271642610068X

Downloaded from https://www.cambridge.org/core. 19 Nov 2026 at 21:32:44, subject to the Cambridge Core terms of use.