Online Comprehension of Verbal Number Morphology in Children with Developmental Language Disorder: An Eye-Tracking Study

Spyros Christou[1], *Carmen Julia Coloma*[2, 3], *Llorenç Andreu*[4], *Ernesto Guerra*[3], *Claudia Araya*[2,5], *Javier Rodriguez-Ferreiro*[1], *Mónica Sanz-Torrent*[1]

1. Faculty of Psychology, Universitat de Barcelona, Barcelona, Spain.

2. Faculty of Medicine, Universidad de Chile, Santiago de Chile, Chile.

3. Center for Advanced Research in Education, Institute of Education, Universidad de Chile, Santiago de Chile, Chile

4. NeuroDevelop eHealth Lab, eHealth Center, Universitat Oberta de Catalunya, Barcelona, Spain.

5. Faculty of Pedagogy, Universidad Academia de Humanismo Cristiano, Santiago de Chile, Chile.

Conflict of interest

The manuscript contains original data which has not been published in any journal, and has not been submitted for consideration elsewhere. The authors have no competing interests to declare.

Funding

This research was supported by Grant EDU2013-44678-P from the Ministry of Economy and Competitiveness of Spain (S.C., M.S.T, L.A., J.R.F.) and the Chilean National Commission for Scientific and Technological Research (CONICYT). Funding from PIA/CONICYT Basal Funds for Centers of Excellence, Project FB0003 (C.J.C., E.G.), and FONDECYT individual grant and 11160657 (C.A.) is also gratefully acknowledged.

Corresponding author

Spyros Christou, Ph.D.

Faculty of Psychology, Universitat de Barcelona

Passeig de la Vall d´ Hebron, 171, 08035, Barcelona, Spain.

Tel: +34933125160

Email: s.christou@ub.edu

Abstract

Purpose: Previous studies have raised the possibility of preserved language comprehension in children with Developmental Language Disorder (DLD) in online tasks and within simple sentence structures. Consequently, we evaluated the capacity of children with DLD to comprehend verbal number agreement in simple sentence structures (i.e., Verb-Object-Subject and Verb-Subject). **Method:** Using an eye-tracking methodology, we conducted two psycholinguistic experiments with 96 Spanish- and Catalan-speaking participants. The sample was distributed into four groups: 24 children with DLD (age range: 4;6-12;6, average age: 7;8), 24 children with the same chronological age (4;6-12;2, 7;8), 24 children with the same linguistic level (4;6-9;4, 6;8), and 24 university students, as language-experts (18-30, 22;5). **Results:** The experimental data indicate that children with DLD can comprehend verbal number agreement, at least under the present experimental conditions. **Conclusions:** The empirical outcomes suggest that number morphology comprehension by children with DLD might be more typical than what it is generally considered to be.

Keywords: Developmental Language Disorder, language comprehension, verbal number morphology, eye movements, Psycholinguistics.

Developmental Language Disorder (DLD) is a severe and persistent disorder in the acquisition and development of oral language, both expressive and receptive, which is not associated with a medical condition or social difficulty, and which may involve one or more components of language to different degrees (Bishop et al., 2017). The term DLD is relatively new. However, until recently, the term Specific Language Impairment (SLI) was the most prominent. Today, it is generally accepted that children who meet these classic SLI criteria fall under the umbrella of DLD (Bishop et al., 2017), so this last term will be used throughout the paper. Many hypotheses regarding the nature of DLD have focused on the grammatical component, because research findings have shown that morphosyntactic problems in children with DLD across languages are significant (Leonard, 2014). These problems relate to a variety of morphological morphemes, such as articles (e.g., Anderson & Souto, 2005; Auza & Morgan, 2013; Morgan et al., 2013), clitic pronouns (e.g., Jacobson & Schwartz, 2002; Morgan et al., 2013; Restrepo & Gutiérrez-Clellen, 2001), prepositions (e.g., Auza & Morgan, 2013b; Grela et al., 2004; Sanz-Torrent et al., 2008), derivational and verbal morphemes (e.g., Bedore & Leonard, 2001; Bishop, 1997; Morgan et al., 2013). Recently, Castilla-Earls et al. (2020) suggested that clitics and verbs present the best diagnostic accuracy as morphological markers of DLD in monolingual Spanish-speaking children. Grammatical difficulties are evident in Spanish-speaking children with DLD, in terms of both production and comprehension (Mendoza, 2016). In language production, most of the problems that children with DLD face relate to aspects of verbal morphology (Bedore & Leonard, 2001; Bishop, 1997; Leonard, 2014). However, only a few studies have tried to clarify whether the problems found in these children's production also characterise their language comprehension (Bishop & Adams, 1992; Marinis & van der Lely, 2007; Montgomery, 2000a, b, 2002, 2004; Montgomery et al., 2017).

Production impairments in children with DLD

Studies in the production of morphology concur that the language of children with DLD shows a higher incidence of incorrect morphemes than that of control groups (Bedore & Leonard, 2001; Morgan et al., 2013). In English, difficulties in the use of the third-person singular in the present tense and the past tense morphemes of regular verbs have been reported (Norbury et al., 2002 Conti-Ramsden et al., 2001; Hoover et al., 2012; Leonard et al., 1997). In Spanish, studies in the production of tense marking show that children with DLD tend to perform more poorly than their age-matched peers in the use of past and present tenses in controlled tasks and that they show a similar pattern to a control group matched by linguistic age. Specific problems are found in the present third-person plural and the past third-person singular and plural (Bedore & Leonard, 2001), as well as in the subjunctive mood (Morgan et al., 2013). However, findings are not consistent with regard to the observation of verbal morphology in spontaneous speech: Bedore and Leonard (2005) reported no differences among children with DLD, age-matched peers, and linguistic controls, while other studies report that children with DLD make more errors than their age-matched peers in tense marking (Grinstead et al., 2014) and that they are also challenged by the use of inflections of mood, time, and person, particularly when irregular verbs are involved in conversation (Buiza et al., 2016). Regarding the production of Spanish verbal morphemes of number, children with DLD, in a longitudinal study (Sanz-Torrent et al., 2008), were observed to produce singular forms when plural forms were required. Overall, Spanish-speaking children with DLD seem to be less troubled by verbal inflections than their English-speaking peers (Guasti, 2017), a conclusion which is in line with the view that difficulties with morphology in children with DLD vary according to the characteristics of the specific languages they speak (Auza, 2009). According to the morphological richness account, built on evidence from crosslinguistic studies of children with DLD (Bedore & Leonard, 2001, 2005; Leonard, 2014; Lidner & Johnston, 1992; Thordardottir, 2008), children with DLD who acquire a morphologically rich language, such as Spanish, for example, will use grammatical morphemes more accurately than children with DLD who acquire a language with less inflectional morphology. Consequently, Spanish-speaking children with DLD may display a greater capacity to use morphemes compared to other children, who speak languages that are less rich morphologically, such as English, for example (Mendoza, 2016).

Comprehension impairments in children with DLD

Research into linguistic comprehension abilities has been scarcer than into expressive linguistic abilities in children with DLD (Muñoz et al., 2014). However, regarding grammatical comprehension, there is a wide consensus that children with DLD have difficulties in this area (Leonard, 2014). This might be explained by the Surface Hypothesis (Leonard 1989, 2014), which proposes that children with DLD show a greater difficulty when processing grammatical elements that are shorter in duration and less salient phonologically. On the other hand, such difficulties could also be explained by limitations in the working memory and the processing speed of incoming input which may affect morphology and syntax (Mainela-Arnold et al., 2008; Muñoz et al., 2014; Weismer et al., 1999). In this respect, children who have a low memory capacity might not present the ability to process and comprehend great amounts of linguistic material (Magimairaj & Montgomery, 2012). Consequently, the grammatical comprehension problems of children with DLD will vary according to the size of the information load they have to hold in their working memory to process a sentence. Thus, the processing of longer sentences leads to greater comprehension difficulties in children with DLD than for those in control groups (Montgomery, 1995). Specifically, the grammatical comprehension of children with DLD may be affected by a reduced short-term phonological memory and also by difficulties in the processing speed required to successfully manage the information load associated with specific tasks, for example, additional visual processing of graphic stimuli (Montgomery, 2000a, b, 2002, 2004; Montgomery & Windsor, 2007). On the other hand, when both length of the sentences, and sentence complexity are simultaneously controlled, it is the latter which seems to affect processing in children with DLD (Muñoz et al., 2014).

Children with receptive problems have difficulties comprehending sentences when non-canonical word order is used (Bishop, 1979; van der Lely & Dewart, 1986; van der Lely & Harris, 1990; Muñoz et al., 2014; Montgomery et al., 2017). In fact, in sentence comprehension tasks where children with DLD must pay attention to word order cues, they respond in a random way (van der Lely & Dewart, 1986; van der Lely & Harris, 1990). Montgomery et al. (2017) found that in the absence of semantic-pragmatic cues, their comprehension of canonical and especially non-canonical sentences is limited. In sum, children with DLD seem to display a higher usage of incorrect morphemes than control groups in linguistic production (Bedore & Leonard, 2001, 2005; Buiza et al., 2016; Grinstead et al., 2014; Morgan et al., 2013; Norbury et al., 2002). Nevertheless, results in comprehension vary depending on the methodology used in its study. Thus, offline methodologies show a more affected comprehension, while online methodologies indicate a more preserved comprehension. In more detail, most studies that addressed language comprehension in children with DLD have been based on offline methodologies. Trueswell (2008) suggested that the eye-tracking methodology, as an online comprehension tool, could be more suitable in examining language comprehension, especially during development. In addition, Andreu et al. (2013) proposed that the previously mentioned offline studies might present limitations in capturing the real linguistic comprehension capacity of children with DLD.

In the present study we aim to evaluate the processing of language comprehension of verbal inflections of number in children with DLD. More specifically, we seek to evaluate the capacity of children with DLD to differentiate between singular and plural verbal inflections in simple sentence structures in the present tense, through the use of eye-tracking technology. By "simple sentences", we refer to those which contain a single predication and no other subordinate sentences, based on traditional syntactic dependency criterion (Real Academia Española, 2010). These simple sentences contain only two or three functional grammatical elements so that we can assess two different basic difficulty levels. In the first experiment, we attempt to evaluate the level of comprehension of verbal morphology of number in sentences containing a subject, a verb and a direct object. In the second experiment, we seek to examine comprehension of an even simpler sentence structure, using transitive and intransitive verbs without direct objects. We argue that the number of elements contained in the stimuli contributes to the difficulty level of each experiment and that this difference should be reflected in the results. Additionally, these sentences were designed following a non-canonical –yet grammatical–structure in Spanish, which located the verb at the very start of the sentence, so that it would be easier for us to assess the comprehension of the verbal number morpheme isolated from the number information contained in the subject. Thus, the experimental tasks were designed using pictures and the auditory input of simple sentences in non-canonical order to evaluate verbal morphemes of number with the smallest possible amount of difficulty and cognitive load. If we accept, as the main hypothesis, that it is characteristics of the verb that guide comprehension of a sentence, then limitations in the processing of inflectional morphology of number will be reflected in eye movements, and consequently, in the comprehension level in the experimental tasks. Accordingly, under the proposed hypothesis, it would be expected that children with DLD will obtain worse results than the control groups. If significant differences appear, the possibility of the existence of a deficit in the comprehension of the grammatical category of number could be sustained, and therefore, a more limited comprehension of verbal language in general. Also, we expect that the experimental task with more linguistic components (Experiment 1) will be harder for children with DLD to comprehend, than the task with fewer components (Experiment 2).

Experiment 1

Method

Participants

Our sample consisted of 96 participants, divided into one experimental group of 24 children with DLD (average age 7;8), two child control groups of 24 participants each, matched by the children's chronological age (Age control group, average age 7;8), by the mean length utterance produced by children (MLU-w control group, average age 6;8), and an adults control group (average age 22;5). Regarding the DLD group, children were equivalent in age to their peers in the Age control group (same year ±2 months). Additionally, each child with DLD was paired with a child from the MLU-w control group, according to the MLU-w calculated in words (±0.5 words). The presence of the adult group, as a language expert control group, provides a stronger experimental validity and adequacy to the stimuli. The contrast between the adult group and the DLD group supports an important comparison reference of the magnitude of the differences between the DLD group and the Age and MLU-w control groups. The three child groups were composed of 7 girls and 17 boys. An extensive evaluation was conducted among

260 children within an age range of 3;9-12;6 years old, in order to form the two child control groups. The children selected were those with the most ideal characteristics and scores in chronological-linguistic terms in relation to the children with DLD. It is important to underline that children with DLD do not compose a monolithic group (Laws & Bishop, 2003). Therefore, a split point was introduced inside the DLD group at 8 years and 3 months. This split point enabled there to be the same number of participants in each subgroup (n=12), and this was more appropriate methodologically, since the youngest child with DLD was 4;6 years old and the oldest 12;6. To go into more detail, two chronological subgroups were created, one made up of younger children (DLD1: n=12, average age 6;0) and another of older children (DLD2: n=12, average age 9;7). The same classification was applied to the Age control group (AGE1: average age 6;3 and AGE2: average age 9;4) and MLU-w control group (MLU-w1: average age 5;4 and MLU-w2: average age 8;2). All participants were native Spanish-Catalan bilinguals, and based on the parental report, all children were classified as bilingual since birth. As reported by Alarcón and Garzón (2011), children in Barcelona are equally efficient in both Spanish and Catalan. However, the presence of Spanish is more widespread. According to the last survey published by the Spanish Ministry of Culture, at the Statistical Institute of Catalonia (Idescat, 2018), with regard to comprehension, 99.8% of the population understands Spanish, and 94.4% understands Catalan (for more information about Spanish-Catalan bilingualism and DLD, see Sanz-Torrent et al., 2008). All three groups of children underwent the same assessment based on different standardized tests: Peabody Picture Vocabulary Test-Third Edition (PPVT-III; Dunn & Dunn, 1997), the Kaufman Brief Intelligence Test (KBIT, Spanish version; Kaufman & Kaufman, 2004), Comprehension Test of Grammatical Structures (CEG; Mendoza et al., 2006), which is the Spanish adaptation of the TROG (Test for Reception of Grammar; Bishop, 1983).

We recruited children with DLD from three different institutions: 1) UTAE (Unidad de Trastornos del Aprendizaje Escolar/Hospital Sant Joan de Déu; Unit of School Learning Disorders/Hospital of Sant Joan de Déu); 2) CREDA Narcís Masó (Centro de Recursos Educativos para Deficiencias Auditivas; Centre of Educational Resources for Hearing Impairments); 3) ATELCA (Asociación del Trastorno Específico del Lenguaje de Cataluña; Association of Specific Language Impairment of Catalonia). Standard diagnostics criterion for SLI (Leonard, 2014; Stark & Tallal, 1981) was used as inclusion-exclusion criteria for the creation of the child samples. Nevertheless, as we stated above, children who meet these classic SLI criteria fall under the umbrella of DLD (Bishop et al., 2017). To confirm the diagnosis of children with DLD we analyzed language samples using the Spanish protocol for the Evaluation of Language Delay (AREL; Pérez & Serra, 1998). Then, using transcripts of spontaneous speech we created language profiles that provided data regarding the children's morphosyntactic skills in language production. Based on the mentioned data, we verify that children with DLD presented a delay of at least 1 year (Bishop, 1997). Additionally, we asked parents and caretakers to give an anamnesis in which they reported functional limitations in academic terms, as well as in socio-emotional terms. We used this information to confirm that children with DLD presented no symptoms of impaired reciprocal social interaction. Also, the written consent of parents and adult participants was obtained and the study was approved by the Ethics Committee of the Universitat Oberta de Catalunya. Finally, in two cases, a girl (age: 6;3) and a boy (age: 6;8) were incapable of following the instructions and fulfilling the experimental tasks and so they were not included in the study. Descriptive data of the groups are summarized in Table 1.

[Please insert Table 1 here]

Materials and Experimental design

We created 32 Spanish sentences with two versions each; one with a singular verbal inflection and another one with a plural verbal inflection (see examples 1 and 2 below). Sentences followed a non-canonical, yet grammatical, sentence structure in Spanish (i.e., verb - object - subject). Regarding this kind of sentence structure, it is important to emphasize that Spanish is a free-word-order **language**. Thus, this dislocation of elements in syntactic structure creates a non-canonical, but still grammatically correct sentence (Real Academia Española, 2010). The idea behind this methodological decision was to assess the comprehension of the morphological information conveyed by the number morpheme in the verb (which was located at the start of each sentence) without the influence of the information carried in the subject of the sentence, which was placed in the third position after the direct object (Figure 1). Each of these sentences was paired with a visual display containing four images. On the one hand, two of these images acted as the target and its corresponding competitor, and on the other hand, the other two acted as distractors. The selection of the target in each sentence depended on whether the verb was inflected in the singular or plural form. Firstly, the chosen verbs for the 32 sentences were selected based on their word length. All verbs were 4 to 9 letter words and the majority had 5 (9/32), 6 (7/32) or 7 (8/32) letters. All verbs were high-frequency words and they were contrasted in a frequency database (Pérez et al., 2003). The least frequent verbs (e.g., 1: Spanish: “plantar”; English: "plant", 2: Spanish: “perseguir”; English: "chase", 3: Spanish: “barrer”; English: "sweep") were cross-checked in terms of familiarity and concreteness. Furthermore, all verbs were chosen based on their simplicity and imageability. Secondly, regarding the target and the competitor, all subjects were also high-frequency words and the majority were people (e.g., 1: Spanish: “la niña”; English: "the girl", 2: Spanish: “los abuelos”; English: "the grandfathers"), which were counterbalanced in terms of gender and number. To create a more user-friendly task for children, a smaller number of subjects were animals (e.g., 1: Spanish: “el caballo”; English: "the horse", 2: Spanish: “las tortugas”; English: "the turtles", etc.). In more detail, the proportion of stimuli with animals was 12.5%. In other words, in every eight stimuli, one was with an animal. Thirdly, while in all cases, the target and competitor were animated subjects, all the distractors were familiar inanimate objects (e.g., 1: Spanish: “el cepillo”; English: "the toothbrush", 2: Spanish: “las almohadas”; English: "the pillows"). The images that acted as distractors were also counterbalanced in terms of gender and number. Finally, all the distractors were chosen so as to avoid any semantic relation they may have to the verb in each sentence. For example, in the stimulus presented in Figure 1, “Bebe agua el caballo” (Literal translation: "Drinks water the horse"; Corresponding English translation: "The horse drinks water"), the distractors (i.e., "the toothbrush" and "the pillows") had no semantic relation to the verb "drink". The visual display of the mentioned example is presented in Figure 1.

(1) “*Bebe agua el caballo*” (Literal translation: "Drinks water the horse"; Corresponding English translation: "*The horse drinks water*")

(2) “*Beben agua los caballos*” (Literal translation: "Drink water the horses"; Corresponding English translation: "*The horses drink water*")

[Please insert Figure 1 here]

We created two experimental lists using a 2 x 1 Latin square by crossing items with sentence type (i.e., singular vs. plural). Consequently, items that appeared in their singular version in one list appeared in their plural form in the other list and the same distractors were used across the two lists. Overall, every item appeared once in each experimental list, and every list presented an equal number of singular and plural sentences (16 sentences each). The auditory stimuli were recorded by a female native Spanish-Catalan speaker at a normal speaking rate of 44,100 Hz. Recordings were edited using a digital audio editor by which we adjusted each sentence with a distance of 1000 ms between each onset (time-windows, see Appendix A), and the average duration per word was approx. 500 ms. The auditory stimuli sounded natural and unedited to adult native speakers and the experimental structure facilitated the subsequent analysis of data. The appropriateness of the stimuli was evaluated by different judges (collaborators and authors of the research) and the selection was based on exclusion criteria that sought the highest possible adequacy and pertinence (regular verbs, clear images, distractors with no semantic connection to the verbs, etc.). For detailed information regarding the experimental sentences see Appendix B.

Apparatus

The experiment was conducted using a Tobii T120 Eye Tracker (with a 120 Hz sample rate) and an integrated 17" TFT monitor set to 1024 × 768 pixels. The visual stimuli were all images of 800 × 600 pixels, presented in video format. Each image consisted of four pictures, each located in the center of one of the four screen quadrants. The four pictures were framed within a rectangle. All the images had a white background, and two black lines, one vertical and one horizontal, were used to divide the four quadrants. The tracker uses infrared LEDs to illuminate participants' eyes and has a spatial accuracy of 0.5 degrees. Stimuli were presented and data were collected through the Tobii Studio Software version 3.3.

Procedure

Stimuli appeared on the monitor of a Tobii T120 Eye Tracker, at a horizontal distance of approximately 60 cm from the eyes of the participant. Before the initiation of the task, a nine-point calibration was performed at the start of the experiment to validate the tracking and the registration of the participants' eye movements. Every point appeared in an image with a white background and rested for about half a second to give the user a chance to focus on it. The Tobii

Studio Software automatically validates the calibration and output of the number of points for recalibration. The experimenter repeated the calibration process whenever this was necessary. After calibration, the participants were presented with four practice trials (two sentences in the singular and two in the plural form) to familiarize them with the experimental task. Each participant was given the following instructions: "You will see some images and you will hear a sentence, search as quickly as possible for the correct image and keep looking at it". The stimuli were presented in random order. Before the appearance of each stimulus, a purple circle appeared always at the centre of the screen for 1000 ms, to draw the participant's gaze. The experiment lasted 6 min. The two tasks were presented during the same session, within a period of 40-45 minutes. First, we presented Experiment 1 and then Experiment 2. A considerable effort was made to create a user-friendly setting with periods of play and drawing between the experimental tasks.

Data analysis

We created four regions of interest that corresponded to the location of each of the depicted images on the display (i.e., the target, the competitor and the two distractor objects) that were slightly larger than the images to accommodate non-perfect calibrations (see Figure 1). Subsequently, we extracted a sampling report which provided participants' gaze location on both the horizontal and vertical axes at a sample rate of 120 Hz (~8ms intervals). Then, we examined every trial per participant and region of interest on every millisecond using the R Project software (R Core Team, 2021) to individualize the visual attention paid to each of the objects. For each of these regions, we calculated the proportion of fixations in time bins of 50 ms. Subsequently, the log-transformed ratio (Arai et al., 2007) between the target and the competitor was computed for each participant and trial. To obtain the log ratio, the proportion of fixation towards the target plus a constant value (i.e., 1) was divided by the proportion of fixation towards the competitor plus that constant. Thus, positive numbers corresponded to a preference for the target, while negative numbers corresponded to a preference for the competitor on the log ratio scale. For visualization purposes, the log ratio was subsequently aggregated by participants and 95% confidence intervals (adjusted for within-participant designs, Cousineau & O'Brien, 2014) were calculated (Figure 2).

A non-parametric statistical approach, cluster-based permutation analysis, was adopted to contrast log ratios between the DLD group and the control groups, as well as against a zero distribution (no preference for target or competitor, see Barzy et al., 2020; Guerra et al., 2021; Helo et al., 2021), given that log ratio around zero expressed no object preference. Cluster-based permutation analysis (Barr et al., 2014; Kronmüller & Noveck, 2019; Kronmüller et al., 2017) provides a useful tool to control for family-wise error rates in auto-correlated data (such as time series). Concretely, we first identified temporally adjacent effects that have the same direction and clustered them together (Barr et al., 2014). Then, a cluster mass statistic was obtained by aggregating, for each cluster, the largest absolute summed *t*-values on each time step, resulting in a distribution of *t*-values. We subsequently compared these *t*-distributions against a null-effect *t*-distribution. Such null *t*-distributions are obtained by permutating the labels of the group of interest (i.e., groups being contrasted) and are based on 2000 iterations for every 50 ms bins with scrambled labels. Finally, we tested the significance of each cluster by calculating the proportion of the sum of the largest *t*-values in the simulated distributions greater than the sum of the *t*-values obtained for the clusters in our data. Following Chan et al. (2018), we considered that proportions smaller than .025 can be described as significant by a two-tailed test.

We conducted a total of ten different contrasts. We first contrasted our group of interest (i.e., the DLD group) against the other three groups (i.e., Age-control, MLUw-control, and adults), as well as the DLD against a zero distribution (Figure 3). Subsequently, we divided the children's data using an age-median split criterion producing two subgroups (i.e., younger children and older children) for each of the three groups of children. We then compared the DLD younger group against the Age-control younger group and the MLUw-control younger group, as well as against a zero distribution (Figure 3, top panel), and we did the same with the older children subgroups.

To determine whether a time bin exhibited statistically significant differences, we used a mixed-effect linear regression on log ratios with a group (e.g., DLD vs. Age-control) as a fixed effect and random intercepts for participants and items. We used the *lmerTest* R package (Kuznetsova et al., 2017) to obtain *p*-values, and then clustered together adjacent time bins showing at least three consecutive (150ms) reliable differences ($p < .05$). We then created ten null-hypothesis distributions of *t*-values; the first four were produced by pairwise randomly permutating each group (i.e., adults, Age-, and MLUw-control) and the zero-distribution label with the label of a group of interest (i.e., DLD). The next six null distributions follow the same logic: They were created by randomly pairwise scrambling each child age-subgroup label (i.e., AGE1, MLU-w1 and AGE2, MLU-w2) and the zero distribution label with the label of the corresponding subgroup (i.e., DLD1 and DLD2). This procedure erases the relation between groups and data, providing the null-hypothesis *t*-distributions. After these *t*-distributions were obtained, we aggregated *t*-values per cluster per simulation and summed the cluster's largest absolute summed *t*-value per time bin making the cluster analysis more conservative than the sum of all *t*-values per cluster, given that, by definition, longer clusters would have larger *t*-values relative to shorter clusters.

Results

Figure 2 shows the log-transformed difference between the target object and the competitor object over time, averaged by participants and divided by each independent group. The shaded area around solid lines represents the within-subject adjusted 95% confidence intervals. Additionally, the figure visually depicts the clusters identified by our statistical analysis. The largest cluster identified pertains to the difference between the DLD group and the adult group (Observed sum $t = 710.23$, $p < .001$), extending from 1050 ms after sentence onset to the end of the trial. By contrast, only a few short clusters show significant differences between the DLD and the Age-control group (from 2200 ms to 2350 ms, from 2950 ms to 3200 ms, and from 3450 ms to 3800 ms with Observed sum *t*-values of 8.69, 14.72, and 17.51, respectively, all *p*-values < .01), and only a single cluster between the DLD and the MLUw-control group from 2900 ms to 3250 ms (Observed sum $t = 21.68$, $p < .001$). Finally, when we contrasted the DLD group log ratios over time against a zero distribution, we observed three short (from 1250 ms to 1550 ms, from 1650 ms to 1900 ms, and from 2450 ms to 2900 ms with Observed sum *t*-values of 18.96, 15.47, and 31.79, respectively, all *p*-value < .05) and one large cluster starting at 3050 ms after the onset of the sentence until the end of the trial (Observed sum $t = 252.75$, $p < .001$).

[Please insert Figure 2 here]

Figure 3 shows the log-transformed difference between the target object and the competitor object over time, averaged across participants and separated by each subgroup of children. The shaded area around solid lines represents the within-subject adjusted 95%

confidence intervals, while the horizontal bars depict the clusters identified by our statistical analysis. When we compared the younger children subgroups (top panel), we observed no statistical differences between the DLD group and the MLUw-control group. By contrast, a short and late yet significant cluster was identified when contrasting the DLD group against the Age-control group from 4300 ms to 4450 ms (Observed sum $t = 8.49$, $p < .001$). Finally, the group of younger children with DLD exhibited statistically significant differences from the zero distribution in two clusters (from 3400 ms to 3600 ms, and from 4050 ms to 5000 ms with Observed sum *t*-values of 12.6 and 87.47, respectively, all p-values < .01).

[Please insert Figure 3 here]

When we compared the older children subgroups (bottom panel), we observed larger differences between groups, and between the group of interest and the zero distribution, relative to that observed in the group of younger children. The earliest significant cluster reflects differences between the DLD group and the Age-control group (from 1600 ms to 1950 ms after sentence onset, Observed sum $t = 17.47$, $p < .001$). These groups were also different between 2900 ms and 3900 ms (Observed sum $t = 53.17$, $p < .001$). A significant difference between the DLD group and the MLUw-control group also appeared to start at 2900 ms and extended until 3700 ms (Observed sum $t = 40.05$, $p < .001$), which was followed by two further significant clusters from 4350 ms to 4650 ms (Observed sum $t = 15.5$, $p < .001$) and from 4800 to the end of the trial (Observed sum $t = 11.17$, $p < .001$). The group of older children with DLD exhibited a significant preference for the target object between 3400 ms and 3600 ms (Observed sum $t = 12.6$, $p < .001$) and from 4050 ms to the end of the trial (Observed sum $t = 87.13$, $p < .001$).

[Please insert Figure 4 here]

Finally, children exhibit a faster and more robust preference for the critical visual object when the sentence referred to a pair of objects (plural) compared to a single object (Figure 4).

Discussion

The purpose of Experiment 1 was to evaluate Spanish verbal morphology of number in an online language comprehension task in children with DLD. Specifically, Experiment 1 focuses on comprehension of the inflection of number presented in the verb within a non-canonical transitive sentence. The most important results obtained indicate the absence of important differences between the DLD group and the Age control group (Figure 2). In more detail, in the object time-window right after the enunciation of the verb (1000-1999 ms), the DLD group presents a comprehension level comparable to AGE and MLU-w control groups. However, the presence of a direct object (i.e., "water") seems to produce an important difficulty in the DLD group and, as a consequence, their preference for looking at the target decreases (1900-2450 ms). In other words, the distinction between target and competitor decreases, and significant differences appear in relation to AGE and MLU-w control groups. From the noun time-window (3000-3999 ms), the ambiguity between target and competitor begins to disappear, and in the final silence time-window (4000-4999 ms) the DLD group reaches high levels of comprehension, comparable to the Age control group. These findings suggest that the capacity of children with DLD to comprehend verbal morphemes of number is preserved, at least under the present experimental conditions. A more detailed analysis revealed significant differences when we include the chronological age predictor as part of the regression and divided all groups into younger and older children (Figure 3). Younger children with DLD (DLD1) show no significant differences throughout the statistical analysis in comparison to the corresponding subgroup of the Age control group (AGE1) or the MLU-w control group (MLU-w1). In fact, all subgroups of younger children manage to distinguish between target and competitor at the end of the task (4050-5000 ms). By contrast, older children of the DLD group (DLD2) present levels of comprehension that differ significantly compared to AGE (AGE2) and MLU-w (MLU-w2) control groups, where the appearance of a direct object produces a slower reaction time (3000-3500 ms). This could be related to the fact that the processing of longer sentences leads to greater comprehension difficulties in children with DLD (Montgomery, 1995; Montgomery & Windsor, 2007). Nevertheless, from the middle of the noun time-window (3500-3999 ms) until the end of the task, older children with DLD (DLD2) significantly increase their level of comprehension which is comparable to AGE2, and significantly higher than MLU-w2. Regarding preference for the target when the verbal inflection was presented in the singular or plural (Figure 4), all groups present higher comprehension performance with plural morphological markers.

Experiment 2

Method

Participants, materials, design, apparatus, procedure and data analysis

Participants, design, apparatus, and procedures were identical to those described for Experiment 1. By contrast, we constructed 32 further non-canonical but still grammatically correct sentences, each with a singular (3) and a plural (4) version, using intransitive or transitive verbs without a direct object (Appendix C). The chosen verbs for the 32 sentences were high-frequency, 3 to 9 letter words and the majority had 5 (8/32), 6 (11/32) or 7 (5/32) letters. We also created 32 further visual displays (Figure 5) that were presented together with the spoken sentences. Finally, data analysis was also identical to that conducted in Experiment 1, with the only exception that of the extent of the period that was analyzed. Given that the sentences in

Experiment 2 were shorter, we only analyzed three consecutive time-windows of 1000 ms each (thus, 3000 ms period). All other aspects of the data analysis approach remained the same.

(3) "*Escribe la niña*" (Literal translation: "*Writes the girl*"; Corresponding English translation: "*The girl writes*")
(4) "*Escriben las niñas*" (Literal translation: "*Write the girls*"; Corresponding English translation: "*The girls write*")

[Please insert Figure 5 here]

Results

Figure 6 displays the log-transformed difference between the target object and the competitor object over time and the corresponding 95% confidence intervals (by participants and adjusted for within-subject designs), divided by each independent group. As for Experiment 1, the figures present the clusters that yield statistical significance in the form of horizontal bars. We replicated Experiment 1 findings when comparing the DLD group with the adult group, where we observed a large cluster starting at 1100 ms after sentence onset to the end of the trial (Observed sum t = 512.06, $p < .001$). We also observed a large cluster, from 1600 ms after sentence onset until the end of the trial, when comparing the DLD group against the zero distribution (Observed sum t = 289.22, $p < .001$). Furthermore, children with DLD differed from the Age-control group in two clusters, one between 1300 ms and 1650 ms from sentence onset (Observed sum t = 22.39, $p < .001$), and another between 2150 ms and 2800 ms (Observed sum t = 31.2, $p < .001$). Finally, they differed less extensively from the MLUw-control group in an early and short cluster (from 1200 ms to 1450 ms, Observed sum t = 14.9, $p < .001$).

[Please insert Figure 6 here]

Figure 7 shows the log ratios difference between the target object and the competitor object over time, averaged by participants and divided by each independent child subgroup. As in previous plots, in the graphs the shaded area around the solid lines represents within-subject adjusted 95% confidence intervals and the colored horizontal bar represents the extension of identified significant clusters. As can be seen in the top panel, DLD1 evidenced a reliable preference for the target from 2750 ms after sentence onset until the end of the trial (Observed sum $t = 99.66$, $p < .001$). Moreover, when we compared DLD1 with the other two child subgroups (AGE1 and MLU-w1), we observed short-lives 200 ms differences between 1400 ms and 1600 ms (Age-controls vs. DLD, Observed sum $t = 12.43$, $p < .001$), and from 1300 ms to 1500 (MLUw-controls vs. DLD, Observed sum $t = 12.29$, $p < .001$) after sentence onset.

[Please insert Figure 7 here]

Children in DLD2, for their part, show a significant preference for the target object starting at 1550 ms after sentence onset to the end of the trial (Observed sum $t = 440.55$, $p < .001$). When compared with the MLU-w2, we found two significant clusters with the opposite direction: an early cluster from 1300 ms to 1650 ms after sentence onset showed an advantage in terms of target preference for the MLUw-controls (Observed sum $t = 10.31$, $p < .001$), while a late cluster from 3650 to the end of the trial (Observed sum $t = 23.4$, $p < .001$) evidenced a greater preference for the target by the DLD group. Finally, the contrast between AGE2 and DLD2 also shows two clusters, yet larger and with a maintained advantage for the Age-control group (from 1300 ms to 1650 ms, and from 2000 ms to 3650 ms from sentence onset, Observed sum *t*-values equal to 18.78 and 89.38, respectively). Finally, all groups exhibit again a preference for the verb in plural form, compared to the verb in singular form (Figure 8).

[Please insert Figure 8 here]

Discussion

The purpose of this second experiment was to examine the comprehension of verbal morphology of number in an even more simple sentence structure than the one used in Experiment 1. In the design of this second experimental task, we used transitive and intransitive verbs without direct objects, in order to reduce possible difficulties and cognitive load as much as possible. According to the results, the DLD group exhibited significantly more differences in comparison to the Age control group at specific moments of the task. In fact, significant intergroup differences between the DLD group and the Age control group were registered in both analyses. In the first analysis (Figure 6), children with DLD present some difficulty in distinguishing between target and competitor in the first 500 ms of the first silence time-window (1000-1500 ms). However, from the second half of the first silence time-window (1500-1999 ms), the comprehension of the DLD group increases significantly and is comparable to Age and MLU-w control groups. Significant differences regarding the comprehension level of DLD and Age groups also appear in the noun time-window (2000-2999 ms), due to a faster response time of the Age control group. Finally, from the last part of the noun time-window until the end of the task, children with DLD reached a level of comprehension similar to the comprehension of children from the Age control group. In the second analysis (Figure 7), the most important differences between the DLD group and the Age control group appear between the older children (DLD2 and AGE2), while the younger children (DLD1, AGE1 and MLU-w1) show no significant differences throughout the statistical analysis. In general, all younger children present low comprehension until the appearance of the agent in the noun time-window. In contrast, older children with DLD (DLD2), despite some difficulty in the first 500 ms of the first silence time-window (1000-1500 ms), show clear disambiguation between target and competitor, and a level of comprehension comparable (1650-1999 ms) to older children from both control groups

(AGE2 and MLU-w2). In the noun time-window, AGE2 significantly outperforms DLD2, but at the end of the task (3650-4000 ms), DLD2 significantly outperforms MLU-w2 and equates to AGE2. Most probably, the developmental advantage of children with typical language acquisition becomes more evident when the complexity of the task decreased. In this sense, in the easiest task, the Age control group identified the target significantly faster compared with children with DLD. Having said the above, it is important to uunderline that, in general terms, both children (with and without DLD) and adults respond better in this second experiment and present a higher level of comprehension, despite the mentioned significant differences. Finally, regarding the differences between verbs in plural form, compared to the verbs in singular form, as in the first experimental task, all groups perform better when the morphological markers are presented in plural form.

General Discussion

A general overview of our results allows us to suggest that children with DLD present a preserved capacity to comprehend Spanish verbal morphology of number, at least under the present experimental conditions (i.e., tracking of eye movements while listening to simple sentence structures represented through visual stimuli), since the DLD group can clearly distinguish between the correct and incorrect stimuli by the end of both experimental tasks. This finding corresponds with results from previous psycholinguistic research carried out using similar methodological procedures (Andreu et al., 2011; Andreu et al., 2016; Andreu et al., 2013; Christou et al., 2020; Christou et al., 2021; Christou et al., 2022).

In a more detailed account of the results, the most significant differences between the DLD group and the Age control group appear in Experiment 2. Here, the Age control group presents a significantly higher level of comprehension than the DLD group for the major part of the noun time-window (Figure 6: 2150-2800 ms), although in the last temporal window the statistical effects disappear and children with DLD present a similar level of comprehension to the children of the Age control group. In the design of the tasks, the number of functional grammatical elements contained in each stimulus was thought to contribute to the difficulty level of each experiment. In this sense, Experiment 2 (example: “Escribe... la niña”; in English, "Writes… the girl") presented a lower level of difficulty than Experiment 1, since it contains fewer grammatical elements than the sentence structure in Experiment 1 (“Bebe agua... el caballo”; in English, "Drinks water…the horse"). In fact, in children with DLD, the processing of longer sentences is related to greater comprehension difficulties when compared to control groups (Montgomery, 1995; Montgomery & Windsor, 2007) and emphasizes the vulnerability of the verbal working memory of children with DLD (Montgomery, 2004; Montgomery & Evans, 2009; Weismer et al., 1999). This effect can be observed in Experiment 1 (Figure 3: 1600-1950 ms), where the presence of the direct object seems to reduce the comprehension of older children with DLD in comparison to the older children of the Age control group. However, the presence of this third element (direct object) could also have affected the performance of the AGE group in their reaction time, which allowed both groups (DLD and AGE) to show a more homogeneous comprehension pattern in this experiment. Most probably, the linguistic advantage that children with typical language acquisition have over children with DLD becomes more evident in Experiment 2 than in Experiment 1, since the Age control group is able to identify the target much faster than the DLD group in a sentence with only two functional elements (subject and verb). In sum, all groups responded better in Experiment 2 mainly because the second task contained fewer grammatical elements. Thus, when the complexity of the task decreased, the pattern of comprehension was more heterogeneous, which allowed more significant differences to appear. In contrast, in Experiment 1 the difficulty affected all children (with and without DLD) in a more homogeneous way, and less significant differences appeared. Regarding the question of working memory, these results suggest, that in terms of comprehension, children with DLD might not exactly develop in the same way as children with typical language development and that their performance results slightly different in this respect. In this regard, it can be argued that the DLD group responds relatively well, but that the Age control group responds significantly better.

Regarding the differences between the verbal inflections in singular or plural form (Figures 4 and 8), a higher comprehension of plural number markers is presented in both experiments and by all groups, especially in the groups of children (DLD, AGE, MLU-w). This effect may have different linguistic explanations, i.e., firstly, verbs with plural markers can only be interpreted in indicative mood. For example, in Experiment 1: Spanish: "Bebe/beben agua el caballo/los caballos"; English: "Drinks/drinks water the horse/horses", after the appearance of the verb in plural form, the agent is necessarily constructed from the indicative mode, i.e., Spanish: "Beben…ellos/ellas"; English: "Drink…they". Conversely, verbs with singular markers can be interpreted in both indicative and imperative moods, i.e., Spanish: "Bebe…él/ella"; English: "Drinks…he/she" (indicative mood) or Spanish: "Bebe…tú"; English: "Drink…you" (imperative mood). This may cause more confusion, and it might have some weight in the discussed effect. On the other hand, another possible explanation may be linked to the Surface Hypothesis (Leonard, 1989, 2014). According to this hypothesis, children with DLD present a greater difficulty when it comes to processing grammatical elements that are shorter in duration and less salient in phonological terms. In sum, we can observe that, in Spanish, the inflections of the singular verbs ("beb*e*") are less salient and lasting in comparison to the inflections of verbs in plural form ("beb*en*").

As mentioned above, the number of functional grammatical elements in each experiment (subject, verb and direct object, in case of Experiment 1; and subject and verb, in case of Experiment 2) was one of the variables which could have had an impact on the level of comprehension of the DLD group, most noticeable in Experiment 2, although children with DLD reached a similar level of comprehension to the Age control group by the end of both tasks. However, an additional variable must also be taken into account in the discussion of the different paths of comprehension observed between the DLD and the AGE group during the processing of the experimental tasks. The importance of this second variable has to do with the non-canonical display of the sentences in the stimuli, which seem to have an important impact on the performance of the DLD group since children with DLD are known to have difficulties comprehending non-canonical word order structures (Bishop, 1979; van der Lely & Dewart, 1986; van der Lely & Harris, 1990; Muñoz et al., 2014; Montgomery et al., 2017). In both experiments (Figures 3 and 7), younger children in all groups show major difficulties in the comprehension of the sentences up to the enunciation of the subject (Figure 3: 3000-3999ms; Figure 7: 2000-2999 ms). After the disambiguation of the sentences through the auditory stimulus of the subject (Figure 3: 4000-4999ms; Figure 7: 3000-3999 ms), all younger children are able to correctly resolve both tasks. In contrast, older children with and without DLD are able to find the target before the enunciation of the subject in both experiments. For younger children with and without DLD, we hypothesize that the main reason for the mentioned effect is the non-canonical order of the sentences. Regarding the performance of older children with DLD in both experimental tasks, we observe that they show significant difficulties in the first time-window after the enunciation of the verb, which makes them more comparable to younger children with and without DLD than to their chronological and MLU-w peers. Most importantly, in both tasks, older children with DLD present a surprising decrease in their comprehension level after the enunciation of the subject (Figure 3: 3000-3999 ms; Figure 7: 2000-2999 ms). This observation suggests that, not only does the non-canonical order of the sentences play a crucial role in comprehension, but that the addition of a new grammatical element creates a greater impact, especially in the most difficult task (Figure 3). It is important to underline that this same experimental sample of children with DLD has shown a typical level of verbal comprehension in canonical sentence structures (Christou et al., 2022). In this sense, we hypothesize that the main reason for the presented effect is the non-canonical order of the sentences, specifically, when the task has more grammatical elements and is more complex. To summarize, children with DLD could comprehend the experimental tasks, but it took them longer to do so, probably due to the syntactic complexity of the sentences. Additionally, the above-mentioned difficulty might be related to a nonsignificant accumulation of mental workload throughout the tasks which eventually leads to a linguistic processing slowdown. In fact, the difficulty of children with DLD when maintaining morphological markers in short-term memory may produce important comprehension problems (Hsu & Bishop, 2014). Having said the above, we argue that the presence of a dysfunctional comprehension in their everyday use of language might be more related to an accumulation of difficulty in linguistic processing, and less to a lack of understanding of specific morphological markers, such as verbal inflection.

The results of both experiments indicate that, under the present conditions, the linguistic comprehension of children with DLD is more typical than that generally considered in the literature and concurs with results from previous research carried out with similar online methodologies (Andreu et al., 2011; Andreu et al., 2016; Andreu et al., 2013; Christou et al., 2020; Christou et al., 2021; Christou et al., 2022). In terms of intervention, these findings are especially relevant when considering the consequences of the language-based problems that children with DLD experience in their day-to-day linguistic performance. Reinforcement of their functional comprehension in natural contexts would require specific support in their oral linguistic abilities. In this sense, scaffolding strategies based on the use of sentences with increasing grammatical complexity could be beneficial to enhance their comprehension of more complex grammatical structures in different and more demanding linguistic contexts. To conclude, it is worth noting that language comprehension cannot be conceptualized in exclusively linguistic terms (Garrod & Pickering, 2009; Pickering & Garrod, 2013; Vygotsky, 1986). In other words, comprehension cannot be reduced solely to syntactic processing, as meaning is not reached until speech is connected to thought (Vygotsky, 1986). Further research should go beyond the isolated analysis of linguistic components, also taking into account their interplay with associative and attention-driven processes which characterize human thought (Stanovich & West, 2000). The study and interpretation of the mentioned interaction might permit a greater insight into the nature and possible causes of dysfunctional language in general, and DLD, in particular.

Acknowledgments

This study was funded by the Ministry of Economy and Competitiveness of Spain and the Chilean National Commission for Scientific and Technological Research (CONICYT). We express our sincere gratitude to all of the children and adults who participated in the study. Also, we thank UTAE (Unit of School Learning Disorders/Hospital of Sant Joan de Déu), CREDA Narcís Masó (Centre of Educational Resources for Hearing Impairments) and ATELCA

(Association of Specific Language Impairment of Catalonia) for their help in setting up the group of children with DLD. Finally, we thank the school Pla de la Vinyes for hosting this research.

References

Alarcón, A. & Garzón, L. (2011). *Language, Migration and Social Mobility in Catalonia.* Leiden: Brill.

Anderson, R.T., & Souto, S.F. (2005). The use of articles by monolingual Puerto Rican Spanish-speaking children with specific language impairment. *Applied Psycholinguistics*, *26*, 621-647. doi:10.1017.SO142716405050332

Andreu, L., Sanz-Torrent, M., Guardia, J., & MacWhinney, B. (2011). Narrative comprehension and production in children with SLI: An eye movement study. *Clinical Linguistics & Phonetics*, *25,* 767-783. https://doi.org/10.3109/02699206.2011.565542

Andreu, L., Sanz-Torrent, M., & Rodriguez-Ferreiro, J. (2016). Do Children with SLI use verbs to predict arguments and adjuncts: Evidence from eye movements during listening. *Frontiers in Psychology*, *6*, 1917. https://doi.org/10.3389/fpsyg.2015.01917

Andreu, L., Sanz-Torrent, M., & Trueswell, J. (2013). Anticipatory sentence processing in children with SLI: evidence from eye movements during listening. *Applied Psycholinguistics*. *34*, 5-44. https://doi.org/10.1017/S0142716411000592

Arai, M., Van Gompel, R. P., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology, 54,* 218-250. https://doi.org/10.1016/j.cogpsych.2006.07.001

Auza, A. (2009). ¿Qué es el trastorno de lenguaje? Un acercamiento teórico y clínico a su definición. *Lenguaje*, *37,* 365-391. https://doi.org/10.25100/lenguaje.v37i2.4893

Auza, A. B., & Morgan, G. (2013a). El uso del artículo en niños hispanohablantes con trastorno específico del lenguaje. *Revista Chilena de Fonoaudiología*, *12*, 3-20. https://doi.org/10.5354/rcdf.v12i0.29513

Auza, A. B., & Morgan, G. (2013b). Uso de preposiciones en el recuento de una historia. Comparación de niños hispanohablantes con y sin trastorno de lenguaje. *Infancia y aprendizaje*, *36*, 35-49. https://doi.org/10.1174/021037013804826573

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*, 404-413. https://doi.org/10.1037/a0031813

Barzy, M., Black, J., Williams, D., & Ferguson, H. J. (2020). Autistic adults anticipate and integrate meaning based on the speaker's voice: Evidence from eye-tracking and event-related potentials. *Journal of Experimental Psychology: General*, *149*, 1097. https://doi.org/10.1037/xge0000705

Bedore, L., & Leonard, L. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*, 905-924. https://doi.org/10.1044/1092-4388(2001/072)

Bedore, L.M., & Leonard, L.B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics. Special Issue: Psychological and linguistic studies across languages and learners*, *26*, 195-225. https://doi.org/10.1017/S0142716405050149

Bishop, D. V. M. (1979). Comprehension in Developmental Language Disorders. *Developmental Medicine & Child Neurology*, *21*, 225-238. https://doi.org/10.1111/j.1469-8749.1979.tb01605.x

Bishop, D. V. M. (1983). *Test for the reception of grammar*. Chapel Press.

Bishop, D. V. M. (1994). Grammatical errors in specific language impairment: competence or performance limitations? *Applied Psycholinguistics, 15*, 507-550. https://doi.org/10.1017/S0142716400006895

Bishop, D. V. M. (1997). *Uncommon understanding*. London: Psychology Press. https://doi.org/10.4324/9780203381472

Bishop, D. V. M., & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. *Journal of Speech and Hearing Research*, *35*, 119-129. https://doi.org/10.1044/jshr.3501.119

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & The CATALISE-2 Consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*, 1068-1080. https://doi.org/10.1111/jcpp.12721

Buiza, J. J., Rodríguez-Parra, M. J., González-Sánchez, M., & Adrián, J. A. (2016). Specific Language Impairment: Evaluation and detection of differential psycholinguistic markers in phonology and morphosyntax in Spanish-speaking children. *Res Dev Disabil. 58*, 65-82. https://doi.org/10.1016/j.ridd.2016.08.008

Castilla-Earls, A., Auza, A., Pérez-Leroux, A., Fulcher-Rood, K., & Barr, C. (2020). Morphological Errors in Monolingual Spanish-Speaking Children With and Without Developmental Language Disorders. *Language, Speech, and Hearing Services in Schools. 51*, 270-281. https://doi.org/10.1044/2019_LSHSS-19-00022

Chan, A., Yang, W., Chang, F., & Kidd, E. (2018). Four-year-old Cantonese-speaking children's online processing of relative clauses: A permutation analysis. *Journal of Child Language, 45*, 174- 203. https://doi.org/10.1017/S0305000917000198

Christou, S., Guerra, E., Coloma, C. J., Andreu Barrachina, L., Araya, C., Rodriguez-Ferreiro, J., Buj Pereda, M. J., & Sanz-Torrent, M. (2020). Real time comprehension of Spanish articles in children with developmental language disorder: Empirical evidence from eye movements. *Journal of Communication Disorders, 87*, 106027. https://doi.org/10.1016/j.jcomdis.2020.106027

Christou, S., Sanz-Torrent, M., Coloma, C.J., Guerra, E., Araya, C., & Andreu, L. (2021). Real-time comprehension of Spanish prepositions and prepositional locutions in bilingual children with developmental language disorder: A study based on eye-movement evidence. *International Journal of Language & Communication Disorders, 56*, 51–71. https://doi.org/10.1111/1460-6984.12581

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

Coloma, C.J., Maggiolo, M., & Pavez, M.M. (2013). Comprensión de narraciones orales en niños con Trastorno Específico del Lenguaje. *Actualidades en Psicología*, *27*, 129-140. https://doi.org/10.15517/ap.v27i115.9277

Coloma, C.J., Mendoza, E., & Carballo, G. (2017). Desempeño gramatical y narrativo en niños con trastorno específico del lenguaje. *Círculo de Lingüística Aplicada a la Comunicación*, *69*, 67-90. https://doi.org/10.5209/CLAC.55314

Coloma, C.J., & Pavez, M.M. (2017). Narrative Comprehension and Language Skills in Chilean Children with Specific Language Impairment. In A. Auza Benavides, & R.G. Schwartz (Eds.), *Language Development and Disorders in Spanish-speaking Children: Literacy Studies, 14*, 261-287. https://doi.org/10.1007/978-3-319-53646-0_13

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment. *Journal of Child Psychology and Psychiatry*, *42*, 741–748. https://doi.org/10.1111/1469-7610.00770

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

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Services, Inc.

Evans, J. (1996). SLI subgroups: Interaction between discourse constraints and morphosyntactic deficits. *Journal of speech and hearing research*, *39*, 655-60. https://doi.org/10.1044/jshr.3903.655

Garrod, S., & Pickering, M. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science, 1*, 292–304. https://doi.org/10.1111/j.1756-8765.2009.01020.x

Grela, B., Rashiti, L., & Soares, M. (2004). Dative prepositions in children with specific language impairment. *Applied Psycholinguistics*, *25*, 467-480. https://doi.org/10.1017/S0142716404001225

Grinstead, J., Lintz, P., Vega-Mendoza, M., De la Mora, J., Cantú-Sanchez, M. y Flores, B. (2014). Evidence of optional infinitive verbs in the spontaneous speech of Spanish-speaking children with SLI. *Lingua, 140*, 52-66. https://doi.org/10.1016/j.lingua.2013.11.004

Guasti, M., T. (2017). Comparison of two types of omission in child language: Articles and third person direct object clitics. *Lingue e Linguaggio, 16*, 279-304. http://dx.doi.org/10.1418/88243

Guerra, E., Bernotat, J., Carvacho, H., & Bohner, G. (2021). Ladies First: Gender Stereotypes Drive Anticipatory Eye-Movements During Incremental Sentence Interpretation. *Frontiers in Psychology*, 1868. https://doi.org/10.3389/fpsyg.2021.589429

Helo, A., Guerra, E., Coloma, C. J., Reyes, M. A., & Rämä, P. (2021). Objects Shape Activation during Spoken Word Recognition in Preschoolers with Typical and Atypical Language Development: An Eye-tracking Study. *Language Learning and Developmen*t, 1-28. https://doi.org/10.1080/15475441.2021.1977643

Hoover, J. R., Storkel, H. L., & Rice, M. L. (2012). The interface between neighborhood density & optional infinitives: Normal development and specific language impairment. *Journal of Child Language*, *39*, 835–862. https://doi.org/10.1017/S0305000911000365

Jacobson, P. F., & Schwartz, R. G. (2002). Morphology In incipient bilingual Spanish-speaking preschool children with specific language impairment. *Applied Psycholinguistics*, *23*, 23-41. https://doi.org/10.1017/S0142716402000024

Kaufman, A. S., & Kaufman, N. L. (2004). *KBIT: Test Breve de inteligencia de Kaufman* (Agustín Cordero & Isabel Calonge, trad. vers. español). Madrid: TEA Ediciones.

Kronmüller, E., & Noveck, I. (2019). How do addressees exploit conventionalizations? From a negative reference to an ad hoc implicature. *Frontiers in Psychology, 10*, 1-10. https://doi.org/10.3389/fpsyg.2019.01461

Kronmüller, E., Noveck, I., Rivera, N., Jaume-Guazzini, F., & Barr, D. (2017). The positive side of a negative reference: The delay between linguistic processing and common ground. *Royal Society Open Science*, *4,* 160827. https://doi.org/10.1098/rsos.160827

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82,* 1-26. https://doi.org/10.18637/jss.v082.i13

Laws, G., & Bishop, D. V. M. (2003). A comparison of language abilities in adolescents with Down Syndrome and children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *46*, 1324-1339.
https://doi.org/10.1044/1092-4388(2003/103)

Leonard, L. B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics*, *10*, 179-202. https://doi.org/10.1017/S0142716400008511

Leonard, L. (2014). *Specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L., Eyer, J., Bedore, L., & Grela, B. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech and Hearing Research*, *40*, 741–753. https://doi.org/10.1044/jslhr.4004.741

Lidner, K., & Johnston, J. R. (1992). Grammatical morphology in language-impaired children acquiring English or German as their first language: A functional perspective. *Applied Psycholinguistics*, *13*, 115-129. https://doi.org/10.1017/S014271640000552X

Magimairaj, B., M., & Montgomery, J., W. (2012). Children's verbal working memory: Relative importance of storage, general processing speed, and domain-general controlled attention. *Acta Psychologica*, *140*, 196-207. https://doi.org/10.1016/j.actpsy.2012.05.004

Mainela-Arnold, E., Evans, J., & Coady, J. (2008). Lexical representations in children with SLI: Evidence from a frequency manipulated gating task. *Journal of Speech, Language, and Hearing Research*, *51*, 381-393.

Marinis, T., & van der Lely, H. K. (2007). On-line processing of wh-questions in children with G-SLI and typically developing children. *International Journal of Language & Communication Disorders*, *42*, 557-582. https://doi.org/10.1080/13682820601058190

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

Mendoza, E. (2016). *Trastorno Específico del Lenguaje* (*TEL*). *Avances en el estudio de un trastorno invisible. Madrid: Pirámide.*

Montgomery, J. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research, 38,* 187-199. https://doi.org/10.1044/jshr.3801.187

Montgomery, J. (2000a). Relation of working memory to off-line and real-time sentence processing in children with specific language impairment. *Applied Psycholinguistics*, *21*, 117-148. https://doi.org/10.1017/S0142716400001065

Montgomery, J. (2000b). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *43,* 293-308. https://doi.org/10.1044/jslhr.4302.293

Montgomery, J. (2002). Understanding the language difficulties of children with specific language impairments: Does working memory matter? *American Journal of Speech-Language Pathology*, *11*, 77-91. https://doi.org/10.1044/1058-0360(2002/009)

Montgomery, J. (2004). Sentence comprehension in children with specific language impairment: effects of input rate and phonological working memory. *International Journal of Language & Communication Disorders*, *39,* 115-13. https://doi.org/10.1080/13682820310001616985

Montgomery, J., & Evans, J. (2009). Complex Sentence Comprehension and Working Memory in Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *52*, 269-288. https://doi.org/10.1044/1092-4388(2008/07-0116)

Montgomery, J. W., Gillam, R. B., Evans, J. L., & Sergeev, A. V. (2017). "Whatdunit?" Sentence Comprehension Abilities of Children With SLI: Sensitivity to Word Order in Canonical and Noncanonical Structures. *Journal of speech, language, and hearing research: JSLHR*, *60*, 1-16. https://doi.org/10.1044/2017_JSLHR-L-17-0025

Montgomery, J., & Windsor, J. (2007). Examining the language performances of children with and without specific language impairment: Contributions of phonological short-term memory and processing speed. *Journal of Speech, Language, and Hearing Research, 50*, 778-797. https://doi.org/10.1044/1092-4388(2007/054)

Morgan, G., Restrepo, M., & Auza, A. (2013). Comparison of Spanish morphology in monolingual and Spanish-English bilingual children with and without language impairment. *Bilingualism: Language and Cognition*, *16*, 578-596. https://doi.org/10.1017/S1366728912000697

Muñoz, J., Carballo, G., Fresneda, M.D., & Mendoza, E. (2014). Grammatical Comprehension in Spanish-Speaking Children with Specific Language Impairment (SLI). *The Spanish Journal of Psychology*, *17*, E45. https://doi.org/10.1017/sjp.2014.47

Norbury, C., Bishop, D., & Briscoe, J. (2002). Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics, 23*, 247-268. https://doi.org/10.1017/S0142716402002059

Pérez, M. A., Alameda, J. R., & Cuetos, F. (2003). Frecuencia, longitud & vecindad ortográfica de las palabras de 3 a 16 letras del Diccionario de la Lengua Española (RAE, 1992). *Revista Electrónica de Metodología Aplicada, 8*, 1-10.

Pérez, E., & Serra, M. (1998). *Análisis del Retraso del Lenguaje: AREL*. Barcelona: Ariel.

Pickering, M., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences, 36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Real Academia Española. (2010). *Nueva gramática de la lengua española. Manual*. Madrid: Espasa.

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

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech & Hearing Research, 39*, 1239-1257. https://doi.org/10.1044/jshr.3906.1239

Rice, M., Wexler, K., & Redmond, S. (1999). Grammaticality judgments of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 943-61. https://doi.org/10.1044/jslhr.4204.943

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

Serra, M., Sellabona, E., Solé, R., Bel, A., & Aparici, M. (2000). *La adquisición del lenguaje.* Barcelona: Ariel.

Stark, R. E., & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders*. *46*, 114-122. https://doi.org/10.1044/jshd.4602.114

Statistical Institute of Catalonia (Idescat). (2018). *Survey on Language Uses of the Population* (Catalan initials: EULP) [online]. IDESCAT. Available at: https://www.idescat.cat/indicadors/?id=anuals&n=10367&lang=en (accessed 19 April 2022).

Thordardottir, E. (2008) Language-Specific Effects of Task Demands on the Manifestation of Specific Language Impairment: A Comparison of English and Icelandic. *Journal of Speech, Language, and Hearing Research, 51*, 922-937. https://doi.org/10.1044/1092-4388(2008/068)

van der Lely, H., & Dewart, H. (1986). *Sentence comprehension strategies in specifically language impaired children. British Journal of Disorders of Communication, 21*, 291-306. https://doi.org/10.3109/13682828609019843 van der Lely, H., & Harris, M. (1990). Comprehension of Reversible Sentences in Specifically Language-Impaired Children. *Journal of Speech and Hearing Disorders*, *55*, 1, 101-117. https://doi.org/10.1044/jshd.5501.101

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT Press (orig. 1934).

Weismer, S. E., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1249–1260. https://doi.org/10.1044/jslhr.4205.1249

Windsor, J., Milbrath, R., Carney, E., & Rakowski, S. (2001). General slowing in language impairment: Methodological considerations in testing the hypothesis. *Journal of Speech, Language, and Hearing Research, 44*, 446-61. https://doi.org/10.1044/1092-4388(2001/036)

Figure Captions

Table 1. Average individual measures per group and pairwise contrasts (Welch two sample *t*-test, two-tailed)

Figure 1. Visual display example for an experimental trial in Experiment 1. Displays were presented together with either a singular-inflected sentence ("Drinks water the horse") or a plural-inflected sentence ("Drink water the horses")

Figure 2. Mean log-transformed fixation proportion differences between target and competitors by group overtime. Grey areas around solid lines represent the within-subject adjusted 95% confidence intervals

Figure 3. Mean log-transformed fixation proportion differences between target and competitors by children subgroup over time. Grey areas around solid lines represent the within-subject adjusted 95% confidence intervals

Figure 4. Mean log-transformed fixation proportion differences between singular (red) and plural (blue) verbs by experimental group and time-window. Grey areas represent the within-subject adjusted 95% confidence intervals

Figure 5. Visual display example for an experimental trial in Experiment 2. Displays were presented together with either a singular-inflected sentence ("Writes the girl") or a plural-inflected sentence ("Write the girls")

Figure 6. Mean log-transformed fixation proportion differences between target and competitors by group and time-window. Grey areas represent the within-subject adjusted 95% confidence intervals

Figure 7. Mean log-transformed fixation proportion differences between target and competitors by children subgroup over time. Grey areas around solid lines represent the within-subject adjusted 95% confidence intervals

Figure 8. Mean log-transformed fixation proportion differences between singular (red) and plural (blue) verbs by experimental group and time-window. Grey areas represent the within-subject adjusted 95% confidence intervals