Developmental Trajectories and Individual Profiles in Sentence Comprehension in Spanish

Analí R. Taboh (artaboh@uba.ar)[1,2,3], Diego E. Shalom[4,2] & Carolina A. Gattei[5,6]
[1]Laboratorio de Neurociencia, Universidad Torcuato Di Tella
[2]CONICET - Universidad de Buenos Aires, Instituto de Física Interdisciplinaria y Aplicada
[3]Facultad de Filosofía y Letras, Universidad de Buenos Aires
[4]Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires
[5]Escuela de Gobierno, Universidad Torcuato Di Tella
[6]CONICET - Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires

Abstract

Sentence comprehension involves integrating multiple cues varying in salience and reliability across structures. A key question is how children weight these cues across sentence types over development. This study examined how Spanish-speaking children aged 3 to 11 years integrate word order, morphological marking, and syntactic embedding. Using a sentence–picture matching task, we assessed comprehension accuracy for five structures differing in canonicity and complexity. Group-level analyses revealed a clear developmental ordering, with early success for structures with canonical word order, more gradual improvement for structures requiring morphosyntactic cues to override linear order, and persistent difficulty when non-canonical order was combined with syntactic embedding. Profile-based analyses showed substantial individual variability, with distinct patterns of strengths and weaknesses across structures among same-age children. These findings indicate that sentence comprehension development reflects both age-related gains in accuracy and differences in cue weighting across structures, highlighting the value of combining developmental and individual-level approaches.

Keywords: language development; sentence comprehension; syntactic complexity; comprehension profiles; individual variability

Introduction

Spoken sentence comprehension is a core component of language development and a prerequisite for successful communication, learning, and academic achievement. To understand a sentence, listeners must integrate multiple sources of information in real time, including the linear order of constituents, the verb's argument structure, hierarchical syntactic relations, and morphosyntactic cues such as case marking or agreement. These cues jointly constrain thematic role assignment, allowing comprehenders to identify who is doing what to whom (Bornkessel et al., 2005).

Evidence indicates that before age 2, children comprehend simple active sentences that follow the canonical word order of their language. Studies with English-learning infants show that they successfully interpret semantically reversible subject–verb–object sentences (Hirsh-Pasek & Golinkoff, 1996). This early competence suggests that children are sensitive to constituent order and use linear order as a cue to assign thematic roles.

This early reliance on linear order can result in a characteristic developmental pattern: relatively good comprehension when the agent appears first, alongside persistent difficulties otherwise. This pattern has been widely reported in the literature and is commonly referred to as the agent-first bias (Abbot-Smith et al., 2017; Gertner & Fisher, 2012). Such a bias enables efficient comprehension of canonical sentences by allowing rapid role assignment based on linear order. However, this strategy becomes ineffective when sentences depart from the canonical pattern, as in structures with non-canonical word order or altered mappings between position and thematic roles.

Indeed, a robust body of cross-linguistic research on spoken language shows that sentence types in which the agent does not appear in canonical sentence-initial position —such as passives, active sentences with non-canonical word order, and object relative clauses— are typically understood later and remain more error-prone across development. For example, studies of typical development have reported that German- and Hebrew-speaking children comprehend active sentences with object topicalization at or below chance level around age six (Biran & Ruigendijk, 2015; Schipke et al., 2012), and that German-speaking children still comprehend them less accurately than sentences without topicalization around age seven (Dittmar et al., 2008). Persistent difficulties beyond this age have been reported for object relatives in English- and German-speaking children (Montgomery et al., 2016; Ruigendijk & Friedmann, 2017).

These difficulties suggest that sentence comprehension is shaped by multiple sources of complexity, including deviations from canonical linear order, morphosyntactic marking, syntactic movement, and embedding. Importantly, these sources of complexity are not uniformly distributed across sentence types: different structures combine them in distinct ways. This makes comparisons across structures particularly informative for understanding how children rely on different cues during sentence processing and how these cues are weighted across development.

Despite the extensive literature documenting age-related differences in the comprehension of non-canonical structures, establishing when children acquire reliable

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

comprehension of different complex structures, and especially their relative order of acquisition, remains challenging. One reason is that relatively few studies have examined the comprehension of multiple complex sentence types within the same group of children (see Friedmann & Reznick, 2021, for a review). In addition, many studies report results based on broad age ranges (e.g., 3–6 years), which makes it difficult to draw precise developmental conclusions in a domain known to undergo rapid reorganization during early and middle childhood (E. Bates et al., 1984; Schipke et al., 2012). This is particularly relevant given evidence that sentence comprehension continues to develop well into the school years, with improvements in accuracy and processing efficiency up to around 11–12 years of age (Dick et al., 2004; Dotti et al., 2018; Wassenberg et al., 2008).

In addition to age-related change, individual variability has emerged as a key feature of sentence comprehension development. Even within typically developing populations, children of the same age can show markedly different levels of comprehension (Montgomery et al., 2016), suggesting the use of different processing strategies or underlying representations. However, most studies either focus on group-level averages or examine individual variability within a single sentence type at a time. As a result, individual performance is typically assessed separately for each structure, without considering how a given child performs across multiple sentence types. This leaves open the question of whether children show distinct comprehension profiles across structures, for example, performing well on one sentence type but poorly on others, and how such profiles relate to developmental trajectories in sentence comprehension.

The present study addresses these gaps by examining the development of auditory sentence comprehension in a large cross-sectional sample of typically developing, Spanish-speaking children. Our goals were twofold. First, we aimed to characterize age-related changes in comprehension accuracy for various structures varying in canonicity and complexity, in order to trace their developmental trajectories and to identify when performance reaches stable above-chance levels. Second, we explored individual differences in comprehension by identifying distinct performance profiles across structures. By integrating group-level developmental trends with analyses of individual variability, this study seeks to provide a more nuanced picture of how sentence comprehension develops in Spanish and how different sources of syntactic complexity interact over time.

Studying Spanish is particularly valuable for expanding the empirical base beyond Anglocentric biases (Blasi et al., 2022). Unlike English, which relies on fixed word order, Spanish leverages relatively flexible word order, overt morphosyntactic markers—such as differential object marking (*a*-marking) and clitic doubling— and verbal morphology to identify thematic roles. This typological diversity is crucial for understanding how different cues interact during development.

Method

Participants

This study included 217 children aged 3 to 11 years (range: 3;4–12;0 [years;months]; *M* = 7;10, *SD* = 2;6, 123 female), distributed across nine one-year age groups (n = 18-30 per group). Participants were native Spanish speakers from Argentina with no neurological or developmental disorders. They were recruited through kindergartens and primary schools in Buenos Aires.

The study followed the ethical principles established in the Declaration of Helsinki and Argentina's Personal Data Protection Law (No. 25326). It was approved by the Ethics Committee of the Facultad de Filosofía y Letras at Universidad de Buenos Aires (protocol number 017). Written informed consent from the parents/guardians of the children and oral assent from the children were obtained.

Materials and Procedure

A sentence-picture matching task, following the paradigm used by Friedmann & Szterman (2006), was designed for Spanish to assess the comprehension of sentences with five syntactic structures: simple active sentences with canonical subject-verb-object word order (SVO), active sentences with non-canonical object-verb-subject order (OVS), passive voice sentences (PASS), and sentences with subject and object relative clauses (SR and OR, respectively); see examples (1-5). This allowed us to compare a baseline structure (SVO) with increasingly complex configurations involving non-canonical word order, morphological marking (marked in bold), and syntactic embedding.

(1) Active sentence with canonical word order (SVO):
*La abuela besa **a** la princesa.*
The grandmother kiss-3SG ACC the princess
'The grandmother kisses the princess.'

(2) Active sentence with non-canonical word order (OVS):
***A** la abuela la besa la princesa.*
ACC the grandmother CL.FEM kiss-3SG the princess
'The princess kisses the grandmother.'

(3) Passive sentence (PASS):
*La abuela **es** bes**ada por** la princesa.*
the grandmother AUX kiss-PTCP by the princess
'The grandmother is kissed by the princess.'

(4) Sentence with subject relative clause (SR):
*La abuela que besa **a** la princesa vive lejos.*
the grandmother REL kiss-3SG ACC the princess live-3SG far
'The grandmother who kisses the princess lives far away.'

(5) Sentence with object relative clause (OR):
*La abuela **a** la que besa la princesa vive lejos.*
the grandmother ACC REL kiss-3SG the princess live-3SG far
'The grandmother whom the princess kisses lives far away.'

Sixty events were constructed using transitive verbs and

animate nouns. For each event, five sentences (one per syntactic structure) were created, all in the present tense, yielding a total of 300 sentences. Stimuli were distributed across five 60-trial lists (12 items per structure) using a Latin square design. Each child was presented with only one list, with list assignment counterbalanced within each age group. Each event was paired with two pictures depicting reversed thematic role assignments (see Figure 1).

Figure 1: An example for a picture pair used in the experiment. These pictures were presented along with one of the auditory stimuli in examples (1-5).

The task was implemented on a tablet using a game-like format. In each trial, the two illustrations of an event were displayed and a recorded sentence was played; the participant tapped the image that matched the sentence. It included a practice phase with five trials (one for each structure) and then a test phase comprising 60 trials. The practice items featured different verbs and characters from those used in the test phase to minimize potential practice or priming effects. During practice, participants received visual feedback (correct/incorrect), whereas no feedback was provided during the test phase. The location of the correct illustration (left or right) was counterbalanced across lists, as well as across the five syntactic structures within each list. Stimuli presentation order was randomized. Participants' picture selection was recorded and accuracy subsequently calculated.

The digit span test from CELF-IV (Semel et al., 2003) was also administered. Testing was conducted individually in a quiet school setting in 12–15-minute sessions.

Analysis

Accuracy To evaluate the comprehension of structures with varying morphosyntactic complexity across development, we fitted a single generalized linear mixed-effects model with a binomial distribution. The probability of matching the sentence with the correct image was analyzed as a function of the interaction between Structure (five levels: SVO, OVS, PASS, SR, OR) and Age (continuous in months, *z*-scored). SVO was set as the reference level (baseline) to assess the relative difficulty of non-canonical structures. Maternal education was included as a covariate to control for its documented impact on complex syntax acquisition (Vasilyeva et al., 2008). The model included random intercepts for participants and items, and random slopes for Structure by participant.

All analyses were conducted using R software (v4.5.1; R Core Team, 2025) and RStudio (2025.9.0.387; Posit team, 2025). Statistical models were fitted with the *lme4* package (D. Bates et al., 2015). The *lmerTest* package (Kuznetsova et al., 2017) was used to obtain *p*-values and the *emmeans* package (Lenth, 2024) with Tukey adjustment for post-hoc comparisons between complex structures (e.g., SR vs. OR).

Comprehension profiles Given the variability observed in language development and in sentence comprehension in particular, an analysis was conducted in order to explore potential comprehension strategies and identify different performance profiles among participants, considering comprehension across all five sentence structures (SVO, OVS, PASS, SR, and OR). For each participant and structure, accuracy was compared against chance level (50%) using a binomial test, following Ruigendijk & Friedmann (2017). With 12 observations per structure, this test yielded significantly above-chance performance with 10 correct responses (83.3%) or more, and significantly below chance performance with two correct responses (16.7%) or less. Based on these thresholds, participants were classified into hypothesized profiles according to their individual pattern of comprehension across the five structures, following previous studies using a similar approach (Szterman & Friedmann, 2014, 2015).

We hypothesized the following profiles, ordered from most to least linguistic competence:

- No difficulties: accuracy significantly above chance in all structures.
- Difficulties only with OR: accuracy above chance for all structures except OR, where performance was at or below chance. This pattern suggests difficulties only when non-canonical word order and embedding occur together, but not when these sources of complexity appear separately.
- Agent-first strategy + salience: accuracy above chance for SVO, SR, and PASS, but below chance for OVS and OR. This profile suggests reliance on a linear heuristic assigning the agent role to the first noun phrase, unless salient morphology provides cues for thematic role assignment, as is the case in PASS.
- Agent-first strategy: accuracy above chance for SVO and SR, but below chance for OVS, PASS, and OR, indicating reliance on a linear heuristic without benefit from morphological salience.
- Difficulties with non-canonical order and/or embedding: accuracy above chance only for SVO, with chance-level or below-chance performance for all other structures, reflecting difficulties with both non-canonical order and embedding, even when they occur separately.
- At-chance performance: comprehension at chance for all structures, suggesting the absence of a predominant comprehension strategy.

Given the stringent criteria for determining significant

differences from chance, some of these profiles required quite extreme performance (i.e., very high or very low). Participants with comprehension patterns that did not match any of the profiles defined above were classified as "Other". This additional category also captured participants whose patterns resembled one of the profiles but did not meet its thresholds.

To complement our theoretically defined comprehension profiles, we applied a *k*-means clustering algorithm to the individual accuracy scores across the five sentence structures, treating accuracy as a five-dimensional vector. *K*-means (MacQueen, 1967) is a simple unsupervised clustering method that partitions observations into a predefined number of clusters based on the patterns in the data itself. Importantly, this procedure does not take participants' age into account; instead, it classifies them solely according to their performance. This approach allows us to examine whether systematic patterns of comprehension —reflecting underlying strategies— emerge naturally from the data, providing evidence for the profiles defined a priori. The number of clusters was set to $k = 7$, matching the number of hypothesized profiles. Clustering was performed using the kmeans() function. We then compared the resulting clusters with the theoretically defined profiles, examining both the correspondence between classification schemes and the distribution of participants across profiles and clusters.

Results

Accuracy

Mean accuracy by structure and age group is shown in Figure 2. Performance significantly exceeded chance at different ages: age 3 for SVO, age 4 for SR, age 5 for PASS, and age 7 for OVS. In contrast, OR structures were comprehended below chance in younger age groups and remained at chance level in the 9- to 11-year-old groups, although accuracy did increase across these older age groups, reaching 65% in the oldest one. Notably, OR clauses exhibited a distinct, non-linear trajectory, with its lowest point at age 7.

Statistical analysis (Table 1) revealed that all non-canonical structures were significantly more challenging than the SVO baseline, with OR showing the most pronounced deficit. Accuracy significantly improved with Age and Maternal education. Critically, the model revealed significant Structure x Age interactions for PASS and OR. For PASS, a positive interaction reflected a steeper developmental slope (i.e., faster catching-up) relative to SVO. Conversely, the interaction for OR was negative, capturing the initial developmental dip and the significantly different growth rate of this structure compared to the canonical baseline. Beyond the comparison with SVO, post-hoc tests confirmed a significant asymmetry within relative clauses, with OR being more difficult than SR ($z = 19.45$, $p < .001$).

An additional model including working memory (WM, measured with backward digit span; n = 186 because most children aged 3 could not complete the task) identified it as a significant predictor ($\beta = 0.383$, $p < .001$). In this model, after controlling for WM, the OR x Age interaction became non-significant ($p = .52$), revealing a slope similar to SVO, while the OVS x Age interaction reached significance ($\beta = 0.428$, $p = .03$), indicating a steeper catch-up rate.

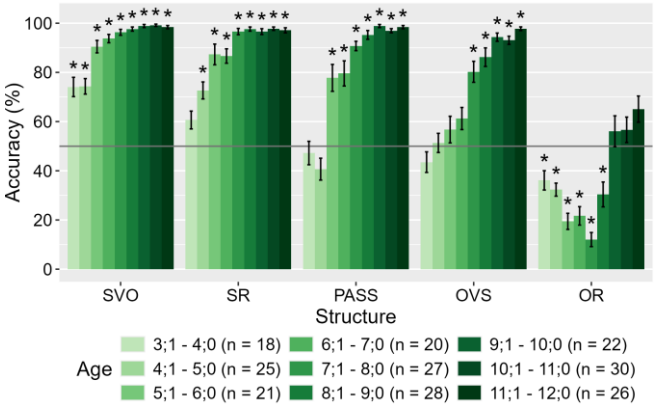

Figure 2: Average accuracy for simple active sentences with canonical word order (SVO), subject relative clauses (SR), active sentences with non-canonical order (OVS), passive sentences (PASS), and object relative clauses (OR) for each age group. Error bars represent the standard error of the mean. The gray line indicates chance level (50%). Asterisks mark performance significantly different from chance after Bonferroni correction ($p < .006$).

Table 1: Model results for comprehension accuracy.

Term	β	*SE*	*z*	*p*
Intercept (SVO)	3.447	0.163	21.19	< .001
Structure: SR	-0.51	0.19	-2.685	.007
Structure: PASS	-0.712	0.206	-3.464	< .001
Structure: OVS	-1.623	0.183	-8.887	< .001
Structure: OR	-4.139	0.199	-20.79	< .001
Age	1.363	0.132	10.29	< .001
Maternal education	0.137	0.052	2.657	.008
SR x Age	-0.191	0.147	-1.3	.194
PASS x Age	0.512	0.176	2.91	.004
OVS x Age	0.102	0.157	0.651	.515
OR x Age	-0.722	0.177	-4.085	< .001

Comprehension profiles

The classification of individual performance showed that the most frequent comprehension profile reflected selective difficulties with OR sentences, with over 40% of the participants in this category (see Figure 3). A smaller group showed good comprehension of all structures (12.9%). Only two participants showed comprehension patterns compatible with the use of an agent-first strategy, with or without benefit from salient morphology. Very few participants exhibited profiles indicating extended difficulty, namely performing above chance only on SVO sentences or at chance across all structures. Finally, close to one third of the

participants exhibited comprehension patterns that did not match any of the theoretically defined profiles. Notably, profiles indicating broader difficulties were only present at early ages, showing a heavily age-dependent distribution.

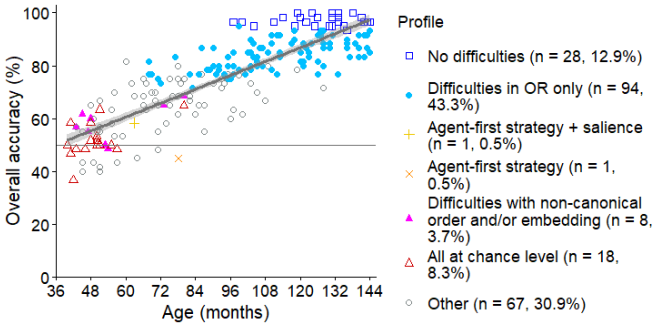

Figure 3: Distribution of children across theoretical comprehension profiles as a function of age. Each point represents one child, and colors and shapes denote the assigned profile. Percentages (rounded to one decimal) and raw counts of children in each profile are indicated in the legend. The gray line indicates chance level (50%), and shading represents the 95% confidence interval around the regression line. OR = object relatives.

To assess whether similar comprehension patterns emerged from the data without defining profiles a priori, we applied *k*-means clustering to participants' accuracy scores for the five sentence types. The comparison between the theoretical profiles and the data-driven clusters is shown in Figure 4. The figure displays mean accuracy for each syntactic structure within each theoretical profile, represented by points connected by solid thick lines, with solid thin lines indicating the range of observed scores. For ease of visual comparison, data-driven clusters are superimposed onto the theoretical profile with the most similar mean accuracy pattern across structures, and are shown using dashed lines for both mean accuracy and range.

Note two main differences between the theoretical and empirical classifications. First, among children with good comprehension of all structures except for OR, the empirical classification distinguished between those with chance-level and those with very low performance on this structure, resulting in two clusters resembling the "Difficulties in OR only" profile (note that there are two sets of dashed lines). Second, no cluster resembled the "Other" profile. Participants categorized as "Other" by the theoretical classification did not meet the binomial test criteria, but were generally close to them. Consistent with this, *k*-means clustering placed all of these participants near the clearly defined theoretical profiles. For example, the "agent-first strategy" and "agent-first strategy + salience" clusters included 22 and 21 children, respectively, compared to only one in each of the equivalent theoretical profiles, and the "difficulties with non-canonical order and/or embedding" cluster included 20 children, compared to eight. Indeed, the empirical clusters generally display wider score ranges, reflecting greater variability among participants than is captured by the theoretical profiles. Overall, of the participants classified into a clearly distinguishable profile (i.e., excluding the "Other" profile), 94% (141 out of 150) were assigned to the same profile/cluster by both classifications.

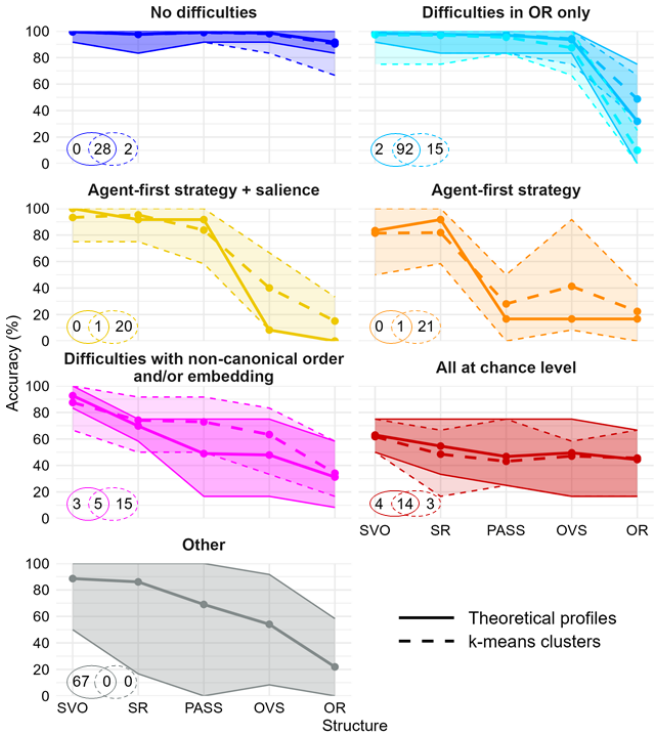

Figure 4: Mean accuracy and score ranges for each syntactic structure within each theoretical profile and data-driven cluster, with both classifications overlaid to facilitate visual comparison. The lower-left corner of each panel indicates the number of participants classified into each profile by the theoretical classification alone (left), into each cluster by *k*-means alone (right), and by both classifications (center). SVO = simple active sentences with canonical word order. SR = subject relative clauses. OVS = active sentences with non-canonical order. PASS = passive sentences. OR = object relative clauses.

Discussion

The present study examined the development of sentence comprehension across a range of syntactic structures differing in canonicity and complexity in Spanish-speaking children aged 3 to 11 years. The results reveal a clear developmental ordering across structures, allowing a more fine-grained characterization of *when* reliable comprehension emerges for each sentence type in Spanish. Canonical SVO sentences were comprehended above chance from the youngest age tested (3 years), with performance reaching near-ceiling levels by age 5. Subject relative clauses showed above-chance comprehension shortly thereafter, at age 4, with performance approaching

90% at age 5. Passive sentences reached reliable above-chance comprehension later, around age 5, and continued to improve gradually across the school years. Active sentences with non-canonical OVS order showed a further delay, with stable above-chance performance emerging around age 7. In contrast, object relative clauses remained particularly challenging, and group-level performance did not reach stable above-chance levels even in the oldest children tested, aged 11 years; however, accuracy did increase in the older groups, suggesting ongoing acquisition of these structures.

This ordering is consistent with the idea that children rely on different sources of information at different stages of development. Early in development, children appear to rely primarily on linear order to assign thematic roles, which supports effective comprehension of SVO sentences and, shortly thereafter, SR, where canonical order is preserved within the embedded clause. With increasing age, children become increasingly able to integrate morphosyntactic cues to override linear order, as reflected in the gradual improvement observed for PASS sentences. In contrast, OVS sentences require the integration of less salient morphosyntactic marking, resulting in later comprehension. When non-canonical order and less salient cues are combined with syntactic embedding, as in OR, comprehension appears to be particularly demanding, leading to a markedly prolonged and unstable developmental trajectory. Notably, this distinct trajectory was statistically accounted for by WM, suggesting that WM capacity underlies the difficulty of object extractions.

Overall, the developmental ordering observed in the present study is broadly consistent with previous findings on sentence comprehension in Spanish. Prior work has documented early mastery of canonical SVO sentences, relatively good comprehension of subject relative clauses from preschool age, and persistent difficulties with structures involving non-canonical word order or complex mappings between syntax and thematic roles, including passive sentences, active sentences with OVS order, and object relative clauses (Muñoz et al., 2014; Pierce, 1992; Presotto & Torregrossa, 2024; Torrens, 2024). However, the age ranges examined and the reported levels of performance vary substantially across studies, making it difficult to establish when stable, above-chance comprehension is achieved for each structure. By testing multiple sentence types within the same group of children and using narrow age bands, the present results refine this picture by showing an early and robust advantage for canonical structures, a more gradual improvement for passives and OVS sentences, and a markedly delayed and unstable trajectory for object relative clauses. In particular, the persistent difficulty observed for object relatives aligns with recent evidence of performance at or below chance in Spanish-speaking children aged 4 to 11 (Llompart & Dąbrowska, 2024; Presotto & Torregrossa, 2024; Zamora et al., 2023).

The observed developmental sequence also bears on the long-standing question of whether acquisition patterns are driven primarily by structural complexity or by distributional properties of the input. The relative ease of structures in which the agent appears sentence-initially may reflect their high frequency in child-directed speech, their lower processing cost, or both. These explanations are difficult to disentangle: frequent structures may be easier to process precisely because they are well entrenched, while structures with lower processing costs may be more likely to occur in the input. Crucially, in the absence of strong cues contradicting an agent-first interpretation, the preference to interpret the first noun phrase as the agent appears to persist into adulthood (Reyes, 2003), suggesting continuity between developmental and mature processing strategies.

Beyond age-related change, the profile-based analyses shed light on the effectiveness and limits of different sentence processing strategies, as reflected in their outcomes across structures. While group-level analyses revealed clear developmental trends, individual profiles showed that children of the same age often relied on distinct combinations of cues across structures. Some profiles were characterized by highly effective processing of agent-first structures, supporting accurate comprehension of SVO sentences and, in some cases, SR. However, the same children often showed marked difficulties with structures that require overriding an agent-first interpretation, such as PASS, OVS, or OR. Other profiles reflected a more flexible but less robust strategy, with intermediate performance across multiple structures. Importantly, these patterns suggest that developmental change does not simply reflect a uniform increase in accuracy, but a gradual reweighting of cues with different levels of effectiveness depending on the structure being processed.

The data-driven *k*-means clustering largely converged with the theoretically defined profiles, while also capturing additional variability not detected by the more stringent theoretical classification. This convergence provides independent support for the existence of qualitatively distinct comprehension strategies and highlights the value of examining performance across multiple structures within the same individuals to understand how efficiency, flexibility, and processing cost interact over development.

Finally, some limitations should be considered when interpreting these findings. Because the study used a cross-sectional design, developmental trajectories are inferred from between-group comparisons rather than directly observed within individuals. Longitudinal studies are needed to track how children transition between comprehension profiles over time and to determine when stable mastery of highly complex structures is achieved. In addition, extending the age range beyond early adolescence would help clarify whether and when group-level performance on OR reaches adult-like levels in Spanish. Future work combining offline accuracy with online processing measures may further complement this approach by providing a more complete picture of sentence comprehension strategies, revealing how different cues are deployed in real time and how online processing relates to developmental change.

References

Abbot-Smith, K., Chang, F., Rowland, C., Ferguson, H., & Pine, J. (2017). Do two and three year old children use an incremental first-NP-as-agent bias to process active transitive and passive sentences?: A permutation analysis. *PLOS ONE*, *12*(10), e0186129. https://doi.org/10.1371/journal.pone.0186129

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A Cross-Linguistic Study of the Development of Sentence Interpretation Strategies. *Child Development*, *55*(2), 341–354. https://doi.org/10.2307/1129947

Biran, M., & Ruigendijk, E. (2015). Do case and gender information assist sentence comprehension and repetition for German- and Hebrew-speaking children? *Lingua*, *164*, 215–238. http://dx.doi.org/10.1016/j.lingua.2015.06.012

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences*, *26*(12), 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, *26*(1), 221–233. https://doi.org/10.1016/j.neuroimage.2005.01.032

Dick, F., Wulfeck, B., Krupa-Kwiatkowski, M., & Bates, E. (2004). The development of complex sentence interpretation in typically developing children compared with children with specific language impairments or early unilateral focal lesions. *Developmental Science*, *7*(3), 360–377. https://doi.org/10.1111/j.1467-7687.2004.00353.x

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German Children's Comprehension of Word Order and Case Marking in Causative Sentences. *Child Development*, *79*(4), 1152–1167. https://doi.org/10.1111/j.1467-8624.2008.01181.x

Dotti, H. M., Corrêa, L. M. S., Rivera, G., Benassi, J., & Formichelli, M. C. (2018). Una evaluación de la comprensión de estructuras sintácticas con alto costo de procesamiento en niños en edad escolar. *Revista Argentina de Ciencias Del Comportamiento*, *10*(2), 37–57.

Friedmann, N., & Reznick, J. (2021). Stages rather than ages in the acquisition of movement structures: Data from sentence repetition and 27696 spontaneous clauses. *Glossa: A Journal of General Linguistics*, *39*(1), Article 1. https://doi.org/10.16995/glossa.5716

Friedmann, N., & Szterman, R. (2006). Syntactic movement in orally trained children with hearing impairment. *Journal of Deaf Studies and Deaf Education*, *11*(1), 56–75. https://doi.org/10.1093/deafed/enj002

Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, *124*(1), 85–94. https://doi.org/10.1016/j.cognition.2012.03.010

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*. MIT Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Lenth, R. V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (Version 1.10.5) [Computer software]. https://CRAN.R-project.org/package=emmeans

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, *1*, 281–297.

Montgomery, J. W., Evans, J. L., Gillam, R. B., Sergeev, A. V., & Finney, M. C. (2016). "Whatdunit?" Developmental changes in children's syntactically based sentence interpretation abilities and sensitivity to word order. *Applied Psycholinguistics*, *37*(6), 1281–1309. https://doi.org/10.1017/S0142716415000570

Posit team. (2025). *RStudio: Integrated Development Environment for R* (Version 2025.9.0.387) [Computer software]. Posit Software, PBC. http://www.posit.co/

R Core Team. (2025). *R: A language and environment for statistical computing* (Version 4.5.1) [Computer software]. R Foundation for Statistical Computing. http://www.R-project.org

Ruigendijk, E., & Friedmann, N. (2017). A Deficit in Movement-Derived Sentences in German-Speaking Hearing-Impaired Children. *Frontiers in Psychology*, *8*, 689. https://doi.org/10.3389/fpsyg.2017.00689

Schipke, C. S., Knoll, L. J., Friederici, A. D., & Oberecker, R. (2012). Preschool children's interpretation of object-initial sentences: Neural correlates of their behavioral performance. *Developmental Science*, *15*(6), 762–774. https://doi.org/10.1111/j.1467-7687.2012.01167.x

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals 4 Spanish edition* (4th ed.). PsychCorp.

Szterman, R., & Friedmann, N. (2014). Relative clause reading in hearing impairment: Different profiles of syntactic impairment. *Frontiers in Psychology*, *5*, 1229. https://doi.org/10.3389/fpsyg.2014.01229

Szterman, R., & Friedmann, N. (2015). Insights into the syntactic deficit of children with hearing impairment from a sentence repetition task. In C. Hamann & E. Ruigendijk (Eds.), *Language Acquisition and Development: Generative Approaches to Language Acquisition 2013* (pp. 492–505). Cambridge Scholars Publishing.

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science*, *11*(1), 84–97. (27894076). https://doi.org/10.1111/j.1467-7687.2007.00656.x

Wassenberg, R., Hurks, P. P. M., Hendriksen, J. G. M., Feron, F. J. M., Meijs, C. J. C., Vles, J. S. H., & Jolles, J. (2008). Age-related improvement in complex language comprehension: Results of a cross-sectional study with 361 children aged 5 to 15. *Journal of Clinical and Experimental Neuropsychology*, *30*(4), 435–448. https://doi.org/10.1080/13803390701523091