

ORIGINAL ARTICLE

# What a thousand children tell us about grammatical complexity and working memory: A cross-sectional analysis on the comprehension of clitics and passives in Italian

Vincenzo Moscati<sup>1</sup> , Andrea Marini<sup>2</sup>  and Nicoletta Biondo<sup>3,4</sup>

<sup>1</sup>Dipartimento di Scienze Sociali, Politiche e Cognitive, University of Siena, Siena, Italy, <sup>2</sup>Dipartimento di Lingue e Letterature, Comunicazione, Formazione e Società, University of Udine, Udine, Italy, <sup>3</sup>Basque Center on Cognition, Brain, and Language (BCBL), Donostia-San Sebastián, Spain and <sup>4</sup>Department of Psychology, University of California Berkeley, Berkeley, CA, USA

**Corresponding author:** Andrea Marini; Email: [andrea.marini@uniud.it](mailto:andrea.marini@uniud.it)

(Received 28 November 2022; revised 21 August 2023; accepted 25 October 2023)

## Abstract

Data from 996 Italian-speaking children were collected and analyzed to assess whether a movement-based notion of grammatical complexity is adequate to capture the developmental trend of clitics and passives in Italian. A second goal of the study was to address the relationship between working memory and syntactic development, exploring the hypothesis that higher digit span values predict better comprehension of complex matrix sentences. The results confirm the validity of a ranking of grammatical structures based on constituent movement, with both clitics and passives developing in parallel and later than canonical SVO sentences. Working memory also shows an effect on sentence comprehension in general, but standard measures (digit span forward/backward) do not show a selective advantage in handling complex constructions such as clitics and passives.

**Keywords:** Working memory; grammatical development; sentence comprehension; syntactic movement

## Introduction

Once young children succeed at producing two-word utterances, there is a steady growth in the array of syntactic constructions that they can understand and produce. Their grammatical competence, however, does not develop homogeneously as various morpho-syntactic structures are mastered at a different pace and a rich literature on language development has focused on the factors of linguistic complexity that affect the emergence of different grammatical structures. Fine-grained comparisons between structure-specific growth trajectories are, however, relatively scarce in language acquisition research for many reasons, such as the practical difficulties of performing

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

longitudinal studies or utilizing samples from large cross-sectional databases of children at different ages. Structure-specific developmental trends are often inferred only indirectly, and even the acquisition trends of grammatical structures that have received the most attention are hard to compare; this is true also for passives and Romance clitics that are the two constructions under investigation in this paper. There is, however, a remarkable source of information that might be exploited: the databases collected for the standardization of language assessment tests. These databases, primarily created for diagnostic reasons, usually cover an array of grammatical structures across different age groups, which are often much larger than the sample of participants tested in conventional experimental settings. Despite the great potential of these databases, their use in studies focusing on grammatical development remains limited. This is mostly due to the fact that participants' performance is considered only in aggregated form and scores collected for the various grammatical constructions are usually collapsed into a single composite score, making a comparison between different structures hard or impossible. This information, however, is available within the databases, and can be extracted through fine-grained analyses that go beyond total scores.

In this study, we looked at a large database of responses from Italian children between 4 and 10 years of age, part of the norming sample used to standardize the *Batteria di Valutazione del Linguaggio* in bambini dai 4 ai 12 anni / Battery for Language Assessment in children aged 4 to 12 years old (BVL 4–12; Marini *et al.* 2015). We targeted specific structures, so as to compare the developmental trend of sentences that differ in the morpho-syntactic realization of the internal argument of transitive verbs.

In Romance languages, including Italian, the internal argument of the verb, the *theme/patient*, appears in post-verbal position in canonical SVO sentences, while it appears preverbally in sentences such as passives and certain clitic constructions. It has been suggested that this phenomenon relates to movement operations, that are in turn associated with different degrees of grammatical complexity in language development. For example, Jakubowicz (2005, 2011) proposed a complexity metric precisely defined in terms of the movement steps involved during the derivation. Her account capitalized on the initial preference for *in situ* wh-questions in French, both in children with typical development (TD) and those with developmental language disorders (DLD). In the same vein, Moscati and Rizzi (2014, 2021) proposed that the developmental sequence of agreement configurations conforms to a movement-based metric of grammatical complexity. Indeed, children between 3 and 5 years of age mastered local agreement configurations in Italian that do not involve movement (e.g., Determiner-Noun agreement) earlier than those characterized by movement (e.g., Clitic-Past Participle agreement).

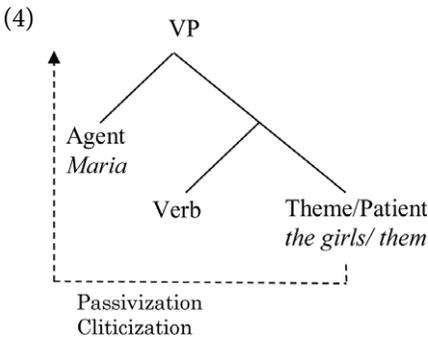
Following these previous approaches, other grammatical constructions, such as passives or clitics, can be investigated and compared once movement is taken as a primitive. In fact, both structures feature the dislocation of the internal argument, as illustrated by the Italian examples in (1–3):

- (1) Maria ha salutato le ragazze (SVO)  
 Maria aux greeted the girls  
 ‘Maria has greeted the girls’

- (2) Maria le ha salutate (Proclisis)  
 Maria them has greeted  
 “*Maria has greeted them*”
- (3) Le ragazze sono salutate da Maria (Passives)  
 the girls aux greeted by Maria  
 “*The girls are greeted by Maria*”

In the SVO sentence in (1), the internal argument *le ragazze* “the girls” occupies its canonical post-verbal position. In (2) and in (3), instead, it surfaces to a position to the left of the finite verb/auxiliary, either as a clitic pronoun or as the subject.

Different approaches exist in the literature to characterize the fine-grained properties of (2) and (3). However, a movement step<sup>1</sup> has been generally assumed in both clitic constructions (see Kayne 1975; Sportiche 1996; Belletti 1999) and passives (Jaeggli 1986, Collins 2005). This is schematically illustrated in (4), where the theme/patient is raised from its post-verbal position to a higher one to the left:



Clitic constructions and passives also share other grammatical properties. For example, they both trigger past-participle agreement in Italian, a property that has also been considered in relation with movement. Although the past-participle does not agree with the object in canonical SVO sentences as in (1), movement triggers past-participle agreement in gender and number in (2) and (3). The uninflected form *salut-ato* in (1) is replaced by *salut-ate* in (2) and (3), a form that carries the same gender and number [–ate = +fem, + plur] of the fronted argument.

Besides considerations based on their grammatical properties in adult language, developmental studies have also unveiled some similarities between the two structures in (2) and (3). Both are produced later than SVO declaratives and are typically mastered only after three years of age. Many factors may be at play here, and they may not be the same for clitics and passives. However, there is at least one family of hypotheses that shares a similar assumption, namely that the observed delay in the acquisition of these structures may be traced back to the movement step in (4). Borer and Wexler (1987) proposed that the comprehension advantage of short passives (e.g., *the window is broken*) over long ones (e.g., *the window is broken by the boy*) stems from the fact that the full movement-based computation is obligatorily required only in the latter case, while short passives may receive a

simpler adjectival reading in which no movement is involved. The asymmetry between long—arguably requiring movement—and short passives has been replicated many times across different languages (Greek: Terzi and Wexler, 2002; Russian: Babyonyshev and Brun, 2003; Portuguese: Rubin, 2009; see also Pierce 1992 for Spanish and the role of post-verbal subjects). In Italian, Volpato *et al.* (2016) investigated children’s comprehension of passives. They manipulated several factors, and, when the comprehension of actional long “*be*” passives was considered, their results indicated that children were able to select the correct interpretation shortly before turning four years old.

This time window roughly overlaps with the age at which children begin to master the syntax of clitics. Italian-speaking children can retrieve the correct antecedent of a clitic computing the local domain by the age of three (McKee, 1992), and this is apparently generalizable across Romance languages (e.g., in French: Hamann *et al.*, 1997; in Spanish, Bauuw *et al.*, 1997). However, other comprehension problems persist longer. For example, Dispaldro *et al.* (2014) analyzed the comprehension of 3rd person object clitics in Italian, testing children’s ability to use gender and number—two features that may be checked through movement—to disambiguate between two possible referents in a picture-matching task. Their results showed that by the age of four, children’s performance was still not adult-like.

More evidence, showing that it is at around the age of four that children’s difficulties with clitics disappear, comes from studies focusing on production. In Italian (Hamann *et al.*, 1997; Tedeschi, 2009), Catalan (Gavarrò *et al.*, 2010), and French (Hamann Rizzi and Frauenfelder 1996; Prévost 2009), the omission of 3<sup>rd</sup> person object clitics disappears at around this age (see also the large comparative study reported in Varlokosta *et al.*, 2016). The acquisitional complexity of clitics cannot be reduced to pronominalization in general, since the production of object pronouns in languages lacking a clitic series—as for example in English—does not seem to pose the same difficulty (Perez-Leroux *et al.*, 2008; Serratrice *et al.*, 2004). Again, as for passives, movement-based accounts have been proposed for the late emergence of cliticization. This hypothesis dates back at least to Wexler (1998), who proposed that the complete sequence of movement steps required by cliticization might be subject to maturation<sup>2</sup>.

More recently, new movement-based accounts have been proposed and object clitic omissions have been linked to a broader class of movement constraints. As shown in Figure 4, the first movement step raises the internal VP argument across the agent. This may give rise to classic intervention effects (Friedmann *et al.*, 2009) that have been documented in both adults and children (for an overview, see Biondo *et al.*, 2022). This line of reasoning was initially proposed by Zesiger *et al.* (2010) for French and more recently by Arosio and Giustolisi (2019) for Italian. According to such accounts, intervention effects are modulated by the degree of similarity, in terms of shared grammatical features, between the moved constituent and the one it crosses over.

This brief review highlights several similarities between the two constructions, both on the theoretical and on the empirical side. Although a unifying account seems beyond reach at the moment, many hypotheses already share the core idea that object movement might be a common source of complexity in language

development. To the best of our knowledge, however, the only explicit comparison between clitics and passives is the production study reported in Manetti (2013). She observed that Italian-speaking children between 3;6 and 4;6 years systematically avoided passives, preferring the use of clitics. Manetti argued that passives could be harder than clitics at this stage, hence the preference for using clitic constructions instead of passives. To date, however, Manetti's work is an isolated case and no comparative study exists in comprehension. Moreover, it is still impossible to have a comparative estimate of the development of the two structures in child language. Our first goal is therefore to enrich the literature in this direction and extend the observation to children between 4 and 10 years old. This is well beyond the time window considered by previous investigations regarding clitics and passives, which have usually focused on the time window between 4 and 6. In most cases, however, children's performance was still not at ceiling, suggesting that some development might still occur in the grammar school years. For this reason, we decided to also include older children.

A second major goal of this study concerns the relationship between working memory and sentence comprehension. As already stated, grammatical development is characterized by great interindividual variability. This may depend not only on sociolinguistic factors (e.g., environmental variables related to parental education and professional profile; Riva et al., 2017; Romeo et al., 2018) but also on a child's cognitive profile. Indeed, morpho-syntactic abilities hinge on the interplay between accumulating linguistic knowledge and growing cognitive skills (i.e., attention, executive functions, and memory; Miyake et al., 2000; Miyake and Friedman, 2012; Marini et al., 2020). Accumulating evidence from studies involving children with TD and DLD suggests that, among the cognitive skills involved in grammatical processing, working memory plays a major role (e.g., Baddeley et al., 1998; Gathercole et al., 1999; Engel de Abreu et al., 2011; Marini et al., 2014; Riva et al., 2017). According to an influential multicomponent model, working memory is a complex system made of different sub-components each dedicated to the storage and management of a specific type of representation (i.e., Visuo-Spatial Sketchpad, Episodic Buffer, Phonological Loop) under the supervision of a Central Executive (Baddeley, 2003, 2007). Together with the Central Executive, the Phonological Loop, made of a passive component (Phonological Short-Term memory, Pho-STM) and an active one (Verbal Working Memory, Verb-WM) which manipulates the information temporarily stored in the Pho-STM, is involved in lexical and grammatical processing. For example, vocabulary growth correlates with Pho-STM abilities as measured through serial span measures (e.g., Digit Forward recall task, Non-word repetition task) (e.g., Baddeley et al., 1998) in children with TD (Gathercole et al., 1999; Leclerc and Majerus, 2010) and DLD (Archibald and Gathercole, 2006). Pho-STM memory and Verb-WM arguably operate synergically in sentence comprehension: Pho-STM holds linguistic material and makes it available for more complex or longer processing cycles; Verb-WM may instead provide the operational resources needed to perform grammatical computations such as those related to syntactic movement. Indeed, long-distance relations generated by the dislocation of a constituent from its original position likely involve working memory, at least until the dependency has been resolved (Gibson, 2000; Lewis and Vasishth, 2005).

To obtain information about the different components involved, serial and composite span tasks can be used: Serial span tasks (e.g., the Digit Forward Recall and the Nonword repetition tasks) have been typically employed to assess the passive capacity of Pho-STM (storage only) whereas composite span tasks (e.g., the Digit Backward Recall and Last Word retrieval tasks) have been related to the functionality of the Central Executive to allocate Verb-WM resources (storage + processing). Results, however, still provide a blurred picture.

In a group of 6-year-old bilingual children, Engel de Abreu *et al.* (2011) found that both Digit Forward and Backward Recall tasks correlated with accuracy in sentence comprehension. However, they measured accuracy considering the overall score obtained by administering the TROG-2 battery (Bishop, 2003). This choice does not allow us to determine whether sentences with an increasing degree of morpho-syntactic complexity also require an increasing amount of working memory resources. This issue was explicitly addressed by Montgomery, Magimairaj, and O'Malley (2008). They looked at *simple* and *complex* sentences. Complex sentences were considered to include different types of non-local relations, such as anaphoras (pronouns, reflexives) and movement relations (passives). Comprehension accuracy was measured for *simple* and *complex* sentences separately. In children between 6 and 12 years old, the authors failed to find significant correlations between measures of Pho-STM and sentence comprehension.<sup>3</sup>

The potential relation between Pho-STM, working memory, and syntactic complexity has been assessed also in studies that focused on clausal embedding. Arosio, Guasti, and Stucchi (2011) analyzed the comprehension of subject and object relative clauses in 9-year-old children. In object relative clauses, the object is moved across the subject, similarly to the movement step shown in (4). This factor has been thought to account for the higher processing load in both children (e.g., Adani, 2011; Arosio *et al.*, 2009) and adults (e.g., Gordon *et al.*, 2001; Grodner and Gibson, 2005; King and Just, 1991; Staub, 2010; Staub *et al.*, 2017; Traxler *et al.*, 2002; Biondo *et al.* 2022), also considering the involvement of memory resources in maintaining active long-distance dependencies (Gibson 2000; Lewis and Vasisht 2005). The results presented in Arosio *et al.* (2011) showed a clear correlation between digit span and object relatives, i.e., the constructions with higher complexity.

The apparent contrast between the findings in Arosio *et al.* (2011) and the ones reported in Montgomery *et al.* (2008) may depend on the different types of sentences that were used in the two studies, thus making any comparison based on grammatical complexity inconsistent. Indeed, in a systematic review of working memory and sentence processing in children, Boyle, Lindell, and Kidd (2013) highlighted that “complexity” has usually been used to refer to either *movement* or *embedding*, two very different kinds of grammatical relations. Boyle, Lindell, and Kidd (2013) treated *movement* or *embedding* as distinct factors and experimentally manipulated them in a sentence comprehension task administered to children aged 4 through 6 years old. They found a significant effect of movement, as sentences with object movement were harder in both moved (i.e., passives) and embedded clauses (i.e., object relatives) with a significant interaction between movement and embedding, meaning that object relatives were the hardest configuration to process. Turning to memory, however, Nonword repetition and Digit Backward Recall tasks, associated with Pho-STM and Verb-WM respectively, did not play any significant

role<sup>4</sup>. More recently, Delage and Frauenfelder (2020) assessed Pho-STM, Verb-WM, and sentence comprehension in children with TD and DLD aged 5 to 14 years old. Notably, the sentence comprehension task included complement clauses as well as subject and object relatives. In children with TD, measures of both Pho-STM and Verb-WM predicted better comprehension only when complex morpho-syntactic constructions were considered. However, again, movement and embedding were not analyzed separately. This limits the results in important ways, and no conclusions could be drawn about the impact of memory resources on object movement.

Taken together, previous studies on sentence comprehension have started unveiling a relationship between language processing and Verb-WM. However, the contribution of the different components of working memory and their involvement with respect to grammatical constructions with different degrees of derivational complexity is still not clear. The mixed pattern that emerges from the previous investigations is arguably related to the selection of the stimuli used in the experiments. Moreover, differences in the age range, in the size of the experimental groups, and in the measures used to assess working memory make it hard to come up with a fully coherent understanding of this complex relationship.

Overall, this study has two main goals: the first is to compare the developmental trend of sentences that do not involve movement (i.e., canonical SVO sentences) with alternative constructions (i.e., Clitic sentences and Passives) that are derived through a dislocation of the internal argument. This is built on a primitive notion of movement-based syntactic complexity, as proposed in Jakubowicz (2005, 2011), Moscati and Rizzi (2014), and Moscati et al. (2020); the second is to investigate the potential relation between Pho-STM and Verb-WM on constructions whose grammatical computation needs temporary storage of non-local dependencies created by a movement operation (i.e., Passives and Clitics).

Cross-sectional data were obtained by administering a grammatical comprehension task to a large cohort of 996 children aged 4 to 10. The participants' accuracy in sentence comprehension was assessed with a picture-matching task, and separate values for each of the three grammatical constructions (i.e., canonical SVO sentences, passives, and clitic clauses) were collected. According to the existing literature, we hypothesized that movement would increase the processing load, making it harder for younger children to understand passives and clitics than SVO sentences. Here, we considered matrix clauses only, so as to exclude the potential confound introduced by clausal embedding. We also aimed to assess the relationship between sentence comprehension and working memory. We, therefore, collected two working memory measures: a serial span measure (i.e., Digit Span forward) that is usually associated with Pho-STM memory and a composite span measure (i.e., Digit Span backward) that is associated with Verb-WM.

## Materials and methods

### *Data availability*

The data and the R analysis code for this study are openly available for download at the following OSF link <https://osf.io/tn23r>.

**Table 1.** Sample size, mean and standard deviation (SD) of children's age across groups

Age group	4	5	6	7	8	9	10	Tot
N	105	189	145	129	134	166	128	996
Mean	4.53	5.44	6.45	7.47	8.50	9.47	10.37	
SD	.28	.29	.29	.29	.28	.29	.28	

### Participants

We analyzed data from a database of 996 children between 4 and 10 years old, taken from the norming sample used in the standardization of the BVL 4–12 (Marini *et al.*, 2015). For descriptive purposes, participants are here grouped according to age and divided into 7 age groups. Age was treated as a continuous variable in all the statistical analyses. Sample size, mean age, and standard deviations for each age group are reported in Table 1.

### Materials

#### *Assessment of working memory skills*

Participants were given the forward and backward digit recall subtests of the Wechsler Scales (1993). The former is a serial span task aimed to evaluate a child's Pho-STM. The latter is a composite span task used to explore the ability of the child to manipulate information in working memory (i.e., the central executive ability to allocate resources to Verb-WM). In the forward digit recall test, children were asked to repeat sequences of digits in the correct serial order. The sequences ranged from 1 to 9 digits that the examiner uttered them at a rate of 1 digit per second. The number of sequences that the child was able to repeat formed the *Forward Digit Recall score*. In the backward digit recall test, children were asked to repeat the sequence of spoken digits in reverse order. The number of sequences that the child was able to adequately invert and repeat formed the *Backward Digit Recall score*.

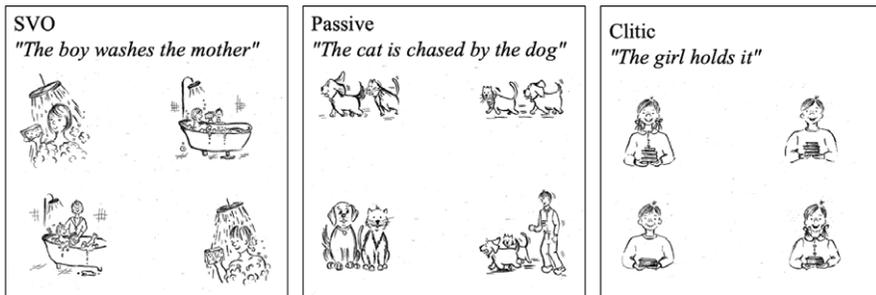
#### *Assessment of grammatical comprehension skills*

All children were administered the 40 items of the *Grammatical comprehension* task of the BVL 4–12. The items belong to different grammatical constructions, and, for each item, children were asked to select the image that accurately represents the meaning of the heard sentence, choosing among four different alternatives. An example is reported in Figure 1.

A correct response for each item was coded as 1, while an incorrect response was coded as 0. Selected constructions were also isolated: we identified matrix clauses with transitive structures, excluding negative sentences. Thirteen items, grouped according to their grammatical structure, were included within this smaller pool. They were divided into SVO, Passives, and Clitic constructions (Table 2). The full list of these selected items is reported in Appendix A.

**Table 2.** Examples of transitive matrix clauses for each grammatical construction

Structure	Movement	N of items	Example
SVO	no	5	il bambino lava la mamma "the boy washes the mother"
Passive	yes	4	il gatto è inseguito dal cane "the cat is chased by the dog"
Clitic	yes	4	la bambina lo tiene "the girl holds it"

**Figure 1.** Examples of the picture-matching task in the 3 conditions. Correct Image: SVO, top-right; Passive, top-right; Clitic, bottom-right.

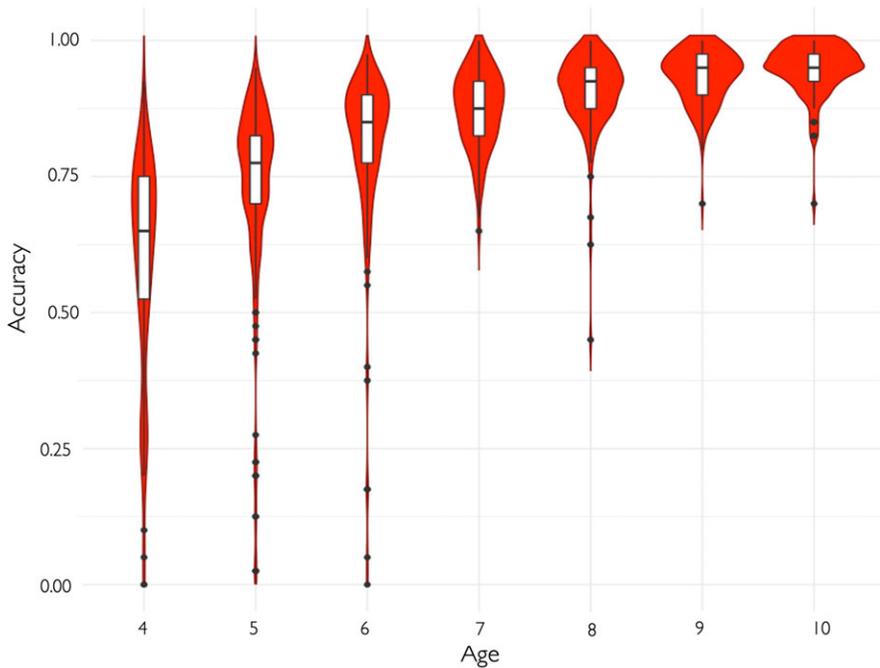
## Results

### **Grammatical comprehension: overall performance**

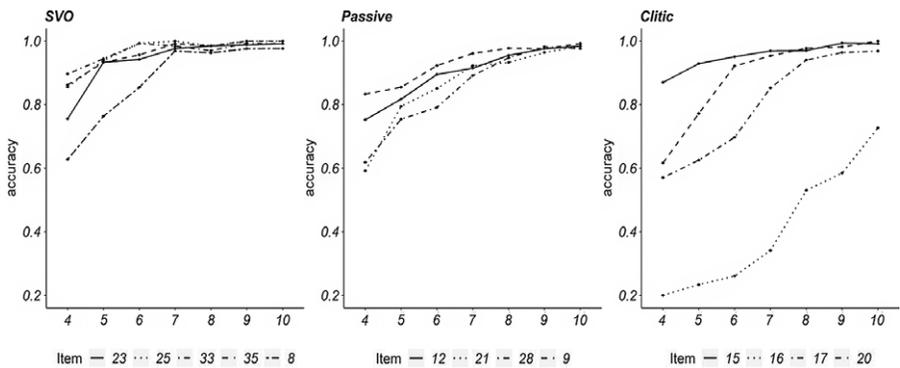
Before considering children's accuracy with respect to the selected constructions, the whole dataset was inspected to determine the overall trend of children's grammatical comprehension across the different ages. Children's aggregated scores were determined by calculating the proportion of correct answers (i.e., *accuracy*) on the total of 40 items. The overall accuracy scores for each age group are reported in Figure 2. As can be seen in this Figure, children's overall accuracy grows constantly with age, with higher variability in the younger groups.

### **Grammatical comprehension: a focus on SVO, Passive, and Clitic sentences**

Before turning to the analysis of the different grammatical structures, the performance on each of the 13 selected items was preliminarily inspected. Figure 3 shows average accuracy values for each item in the SVO, Passive, and Clitic conditions. As can be seen from the plot, item 16 (see Appendix C) in the Clitic condition shows remarkably low accuracy values. Namely, this specific clitic sentence presents difficulties that persist even in older children while all the other items reach ceiling levels by the age of 8. Low accuracy values for item 16 are unlikely to be primarily related to grammatical development. We therefore decided to remove this item from further statistical analyses.<sup>5,6</sup>

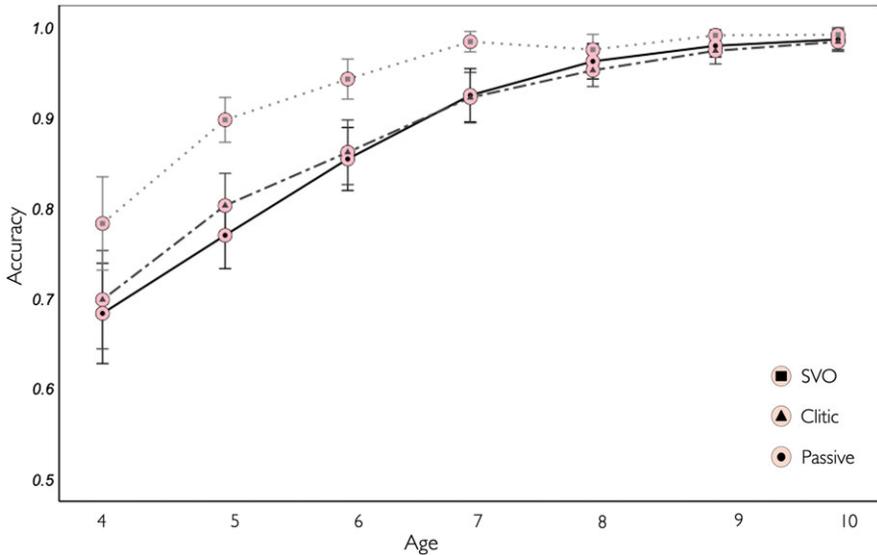


**Figure 2.** Overall accuracy (proportion of correct answers on total) on the Grammatical Comprehension task. Median, interquartile range, and probability density estimation are reported for each age group.



**Figure 3.** Average accuracy values for single items in SVO, Passives, and Clitic structures.

The average proportion of correct responses for the comprehension of grammatical structures without (i.e., SVO) and with (i.e., clitic and passive sentences) movement is reported for each age group in Figure 4. Overall, the three grammatical constructions follow different developmental trends that are clearly



**Figure 4.** Developmental trend across ages and structures: accuracy for SVO, Passives, and Clitics. Error bars indicate  $2 * SE$ . SVO = Canonical Subject - Verb - Object sentence; Clitic = Sentences with a 3rd person object clitic; Passive: Long-passive sentences.

distinct until age 7. The two structures with movement appear to be the most challenging, a result that is in line with the previous literature.

In order to assess the effect of grammatical structure on accuracy, a logit mixed-effect model analysis (Jaeger, 2008) through the package *lme4* (Bates et al., 2014) in R Studio (RStudio Team, 2020) was performed. Compared to traditional non-parametric/parametric tests adopted in previous studies, mixed-effect models have the advantage of accounting for by-subject and by-item variabilities through the definition of a random effect structure. The random effect structure of the model adopted for this study included varying intercepts and slopes for Item and Subject grouping factors. The fixed effects were contrast-coded depending on our hypotheses (Schad et al., 2020; see also Brehm and Alday, 2022), by using the *stats* package (R Core Team, 2021). In particular, the Helmert contrast was adopted. This contrast allowed us to compare the average accuracy of the easier canonical structure (SVO) with the average accuracy of two more challenging grammatical constructions involving movement (i.e., clitic and passive sentences), and to compare the two complex structures with each other. The model thus included the following fixed effects: *Movement* (Passive/Clitic sentences vs. SVO), *Complex structures* (Clitic vs. Passive sentences), centered *Age*, as well as the interaction between age and the two structure-related contrasts. The output of the mixed-effect model analysis is reported in Table 3 and reveals a significant effect of *Age*, *Movement*, and a significant interaction *Movement:Age*. No significant effects were found when comparing accuracy in the comprehension of the two most complex structures (i.e., Passives vs. Clitics).

**Table 3.** Summary of model estimates with Syntactic Movement and Age as predictors, including comparisons between Complex structures. Standard errors, *z*-values, and *p*-values for the accuracy data in the first analysis. Formula of the model: Accuracy  $\sim 1 + (\text{Complex structures} + \text{Movement}) * \text{Age} + (\text{Complex Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

	Estimate	SE	<i>z</i>	<i>p</i>
Age	0.71	0.04	20.17	<0.01
Movement	-0.5	0.13	-3.72	<0.01
Movement:Age	-0.05	0.02	-2.31	0.02
Complex structures(Clitics vs. Passives)	-0.02	0.24	-0.09	0.93
Complex structures(Clitics vs. Passives):Age	0.03	0.03	0.94	0.35

**Table 4.** Sample size, mean and standard deviation (SD) of children's age across groups, in the analysis assessing the role of memory

Age group	4	5	6	7	8	9	10	Tot
N	91	172	141	129	134	166	128	961
Mean	4.52	5.43	6.46	7.47	8.50	9.47	10.37	
SD	.29	.29	.29	.29	.28	.29	.28	

**Table 5.** Correlations between Accuracy and Digit Recall Scores. Results from partial Pearson's correlation analyses controlling for a potential effect of age

Accuracy	Digit Recall	Estimate	Statistic	<i>P</i> -value
SVO	Forward	.10	2.79	.01
	Backward	.14	3.81	<.01
Passives	Forward	.14	3.69	<.01
	Backward	.14	3.67	<.01
Clitics	Forward	.10	2.72	.01
	Backward	.16	4.40	<.01

### **Relation between working memory and sentence comprehension**

The potential relation between working memory and sentence comprehension was addressed by running correlation analyses and linear mixed-effect model analyses. The analyses were conducted on 961 participants (see Table 4) since digit recall scores of 35 participants were missing.

The strength of the relationship between working memory (i.e., forward and backward digit recall scores) and sentence comprehension while controlling for the effect of age was assessed with a series of partial correlations (see Table 5). Such correlations were performed in R through the *ppcor* package and by using the default Pearson method to calculate partial correlation coefficients (Kim, 2015).

**Table 6.** Forward digit recall. Summary of model selection and estimates, standard errors, z-values, and p-values of the best-fitting model for forward memory. Formula of the best-fitting model: Accuracy  $\sim 1 + (\text{Complex structures} + \text{Movement}) * \text{Age} + \text{Forward} + (\text{Complex structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

	Estimate	SE	<i>z</i>	<i>p</i>
Age	.62	.04	<b>15.17</b>	<b>&lt;.01</b>
Forward	.14	.03	4.44	<b>&lt;.01</b>
Movement	-.47	.13	<b>-3.49</b>	<b>.01</b>
Movement:Age	-.05	.02	<b>-2.37</b>	<b>.02</b>
Complex structures(Clitics vs Passives)	.02	.24	.09	.93
Complex structures(Clitics vs Passives):Age	.03	.03	.79	.43

**Table 7.** Backward digit recall. Summary of model selection and estimates, standard errors, z-values, and p-values of the best-fitting model for backward memory. Formula of the best-fitting model: Accuracy  $\sim 1 + (\text{Complex Structures} + \text{Movement}) * \text{Age} + \text{Backward} + (\text{Complex Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

	Estimate	SE	<i>z</i>	<i>p</i>
Age	.55	.05	<b>11.77</b>	<b>&lt;.01</b>
Backward	.22	.04	<b>5.25</b>	<b>&lt;.01</b>
Movement	-.47	.14	<b>-3.49</b>	<b>&lt;.01</b>
Movement:Age	-.05	.02	<b>-2.36</b>	<b>.02</b>
Complex structures(Clitics vs Passives)	.01	.24	.03	.97
Complex structures(Clitics vs Passives):Age	.02	.03	.70	.49

These analyses showed the presence of a positive correlation between comprehension accuracy and both forward and backward digit scores for all three structures supporting the existence of a significant relation between the comprehension of sentences without (i.e., SVO) and with movement (i.e., clitics and passives) and both Pho-STM and Verb-WM.

To further assess the impact of working memory on the acquisition of different syntactic structures, we added forward and backward span values in our mixed-effect model analysis, which was performed by adopting the same packages used for the first analysis. The main effect of memory and its interaction with sentence types was tested in a (forward) stepwise fashion. The models were compared by using *Anovas*. The best-fitting models for forward memory and backward memory are reported in Tables 6 and 7, respectively. The output of the model selection analysis is reported in Appendix B. Table 6 shows that in addition to the main effects of grammatical structure and age and their interactions, the forward digit recall span also significantly increased the probability of providing the correct answer. Results are similar for the backward digit recall span, which significantly increased the probability of providing the correct answer (see Table 7). Children of the same age with higher backward digit recall values and with higher forward digit recall values are seen to provide the correct choice more often than children with lower Backward and/or forward digit recall values.

## Discussion

From the analysis of our database of children between 4 and 10 years old, two major results can be reported. First, the acquisition of grammatical structures progresses along different trajectories, and the presence (or absence) of dislocations predicts their growth curves. Second, sentence comprehension in children is also related to working memory resources, with a generalized effect across the sentence types considered here. We discuss these two results in light of previous developmental studies.

The first finding supports the idea that different grammatical structures are not acquired at the same pace and that movement can provide an explanation for this developmental difference. This hypothesis had already been expressed in slightly different terms in Jakubowicz (2011) and Moscati and Rizzi (2014, 2021) to adhere to the specific features of *wh*-movement in French or agreement in Italian. Overall, the data coming from *wh*-movement, agreement configurations, and direct-object fronting provide a coherent picture, that is a movement-based dimension of grammatical complexity.

In relation to this first result, of note is the fact that the comprehension of clitic and passive sentences did not show a distinct pattern among participants at any age. This trend is shown in Figure 4 and is confirmed by the general mixed-effect model analysis. Due to the nature of our dataset, we cannot exclude that differences indeed may exist between the two constructions and that they can be brought to light by targeted experimental manipulations. Several alternative theoretical notions of movement have been proposed to qualify cliticization or passivation, and endorsing just one or some of them goes beyond the opportunities offered by our dataset. In this sense, datasets used for standardizations provide a formidable resource in terms of sample size and the variety of structures that can be contrasted, not to mention the potential for cross-linguistic comparisons. The tradeoff is that they are designed to evaluate a broad spectrum of linguistic dimensions so that the quest for variety hampers the creation of controlled minimal pairs: the distinctive feature of lab-based experiments. Being aware of the limitations of our tool, for the time being, we take our results at face value and assume that the development of clitics and passives is contingent upon the availability of the general movement step sketched in (4). From this point of view, our study is the first one that explicitly compared and found strong parallelism in the comprehension of clitics and passives.

This can be considered in relation to the only other study that, to the best of our knowledge, has discussed the relationship between clitics and passives in Italian: the production experiment reported in Manetti (2013). This study showed that approximately by four years of age, Italian-speaking children avoid the production of passives in contexts where adults may find them appropriate. In that study, children often used a different answering strategy, using a clitic instead of a passive. Manetti's results show that, *in production*, the clitic option is ranked above the passive one, with the implication that the development of the two structures is asymmetric. Clearly, there are important differences between Manetti's study and ours, leading to more general considerations on the divergence between production and comprehension. For example, production may be more severely affected by performance limitations that may obscure some truly grammatical options. It is

conceivable that the availability of movement unlocks both passivation and cliticization, but that additional limitations in production only apply to passives. For example, issues related to auxiliary selection may delay the production of passives.<sup>7</sup> On the other hand, *comprehension* targets the structures of interest more directly and the constructions under investigation are generally provided in the prompt, already formed. A forced-choice comprehension task might thus reveal substantial parallelism in the grammatical competence of passives and clitics, avoiding complications that would occur downstream in production.<sup>8</sup>

Turning to our second result, we found that the comprehension of sentences either with or without movement (i.e., SVO vs. Clitics and Passives) correlated with measures of both Pho-STM and Verb-WM. We maintained previous assumptions under which serial and composite span tasks (i.e., forward digit recall and backward digit recall, respectively) are the expressions of the working capacities of the two subsystems. Previous results diverge on the conclusions that can be drawn regarding the role of working memory in supporting complex sentence comprehension, and this could be due to at least two reasons. The first, already discussed, is related to the different grammatical structures that fall under the definition of “complex.” Here we adopted an explicit notion of grammatical complexity based on movement. In doing so, sentences with clausal embedding have been filtered out from the dataset. This allowed us to eliminate at least one of the confounds pointed out by Boyle, Lindell and Kidd (2013).

A second reason for the differences among previous studies may be related to the characteristics of the participant samples, including sample size. For example, Montgomery, Magimairaj, and O’Malley (2008) did not find any significant effect of memory measures on sentence comprehension in a sample of 51 monolingual children, scattered across an age interval between 6 and 12 years of age. Engel de Abreu et al. (2011) found instead that working memory had a significant impact on comprehension, looking at a larger group of 109 six-year-old bilingual children. Apart from the important distinction between monolinguals and bilinguals, the number of participants in Engel de Abreu et al.’s study was more than double the size compared to the sample in Montgomery et al. Therefore, we argue that the effects of working memory might go undetected if the study’s statistical power is limited by a low number of participants across wide age ranges as this may lead to high variability (as in Montgomery et al., 2008). In our study, our sample was between 10 and 20 times larger than the ones considered in previous investigations and we found significant working memory effects. We believe that this is one of the main advantages of exploiting large, standardized datasets.

Another open question concerns the non-selectivity of working memory on the comprehension of simple and complex sentences: both simple SVO sentences and Clitics/Passives correlated with higher backward and forward digit recall scores.<sup>9</sup> On one hand, this confirms that the Central Executive and the Phonological loop play a role in sentence comprehension; on the other, it suggests that none of the two subsystems provides resources that are directly involved in the resolution of distance relations. Several models of sentence comprehension, for example the ones in Gibson (2000) and Lewis and Vasisht (2005), have assumed that working memory provides crucial support for the temporary storage of long-distance dependencies. Under this view, we would have expected an interaction between working memory

and clause types. In our opinion, there are at least two possibilities that need to be considered in relation to the lack of interactions.

The first is that passives and clitics are relatively light in terms of working memory requirements so that the effects of working memory would be almost indistinguishable from SVO matrix clauses. If this is correct, in the absence of more sensitive tasks or WM measures, it would be appropriate to increase the memory load and look at more complex sentences, for example, object relative clauses. Here, working memory may be more heavily involved than in subject relatives, due to the greater length of the dependency and the potential insurgence of intervention effects. In fact, Arosio *et al.* (2011) showed that, in Italian 9-year-olds, D-span measures modulate the comprehension of object—but not subject—relatives.

This first line of reasoning is essentially based on the *quantity* of the working memory resources involved. There is, however, a second possibility, based on their *quality*. Memory tasks such as backward and forward digit recall may be only indirectly related to sentence processing: processing relations between constituents is arguably different from repeating sequences of digits. In this respect, our study has shown that the phonological loop and the central executive support language acquisition in general. Reiterating sentences in the Phonological Loop and maintaining them active in the Verbal WM may arguably enlarge their processing window. However, this kind of working memory may not fully overlap with other kinds of mnemonic resources that could more directly support the elaboration of movement dependencies. We must therefore leave open the possibility that other *syntactic* working memory resources would better correlate with grammatical complexities associated with long-distance dependencies.

## Conclusions

By analyzing the data collected through the standardization of grammatical tests and their use as screening tools for language-related developmental disorders, we started exploiting a new data source that is potentially available in many languages. Moving from Italian, we showed that these large datasets provide valuable information for tracing structure-specific developmental trends, as in the comparison between SVO declaratives and constructions with the dislocation of the internal argument, as clitics and passives. In our view, this creates a virtuous cycle, from language development research to clinical practice, and back. Data on syntactic structures that have been initially included in the test batteries due to previous research in language development are later analyzed and considered so to inform current research with new evidence coming from large-scale samplings. Results from our analysis provide new support to movement-based complexity metrics proposed in recent years (Moscati and Rizzi 2014; Moscati *et al.* 2020).

In addition to this, finer-grained analyses that go beyond total scores offer a better vantage point to shed light on the relation between the operations that underlie grammatical development and other cognitive abilities. In line with previous results, our study supports the impact of working memory on comprehension tasks. However, it shows that the kind of working memory resources measured through forward and backward digit recall does not offer a

greater advantage in processing long-distance dependencies created by movement. This opens the possibility that such measures cannot accurately quantify a different kind of working memory—that we can provisionally label *syntactic* working memory—that more directly subserves syntactic computation.

A final concluding remark concerns clinical practice. New analyses per grammatical structure, that go beyond total scores, might provide practitioners with construction-specific reference levels, thus improving the precision of the screening tools. This has important implications for language assessment. The descriptive depth of the test batteries could be enhanced with structure-specific acquisitional trends, adding more refined baselines for the evaluation of possibly distinct subgroups of children with DLDs.

**Replication package.** The data and the R analysis code for this study are openly available for download at the following OSF link <https://osf.io/tn23r>.

**Competing interests.** The authors declare no conflict of interest.

**Ethics statement.** This study has been approved by the ethical committee of the Scientific Institute “IRCCS Eugenio Medea” in conformity with the Declaration of Helsinki.

## Notes

1 A major difference between the different analyses is on the proper characterization of the movement in (4). Analyses vary with respect to the exact specification of the maximal constituent being moved and to the assumption of an additional preliminary movement operation (see Collins 1995 for passives and Belletti 1999 for clitics. For a different analyses, Kayne (1975) and Sportiche (1996/98) for clitics, Jaeggli (1986) for passives.

2 See Moscati and Tedeschi (2009) on Wexler’s Unique Checking Constraint and the development of clitic—past participle agreement in Italian.

3 Only values obtained in a dual-processing paradigm correlated with sentence comprehension. Of note, however, they employed a Processing-Storage Task based on the recall of the final word of the sentence. As such, this measure is based on sentence processing and is itself an indication of grammatical proficiency. For this reason, such a measure may only be a different way to assess sentence-processing abilities, weakening the interpretation of the observed relation between Verb-WM and sentence comprehension skills.

4 Only sentence repetition had an effect. This task however is tightly related to language processing itself. See Boyle, Lindell and Kidd (2013) for a discussion on the Episodic buffer.

5 As a further control, the analyses were rerun also including item 16. Notably, even including this item, the models replicated the main effects and the interactions of the models reported in the next paragraphs.

6 Clitic structures differ in the grammatical specification of the traits of the clitic and the subject. In items 16 and 17, both the subject and the clitic are singular, but they present a gender mismatch. In items 15 and 20, the subject and the clitic share instead the same gender, but they differ in number. Details are provided in Appendix A. We ran a mixed-model contrasting number mismatch items (15,20) against gender mismatch items. Accuracy was significantly higher for number mismatch items. This result is however most likely due to the extremely low performance of Item 16. Across all ages, its mean accuracy is 41%. The average accuracy of all the remaining items with a clitic is instead much higher, between 80% and 95%.

7 Manetti found that passives formed with the auxiliary “venire” emerge earlier than *be*-passives (see also Volpato et al. 2016). Speculating on this, the choice of the auxiliary *be* in Italian, and in many other languages including English, is compatible either with an adjectival or a passive reading. This could be a source of confusion, both in production and in parsing. In connection with this, Grillo et al. (2019) have shown that online processing difficulties with passives vanish in German. In this language, no ambiguity exists between adjectival passives and verbal passives; they are lexically disambiguated by the alternation of two auxiliaries (*sein/werden*).

8 As a matter of fact, production generally lags behind comprehension. This may be true also in the case of clitics. Grütter (2006) showed that at the stage in which children still omit object clitics in French, they are already able to correctly exclude a referential null-object interpretation, equivalent to a silent pronoun, an option not available in their target grammar.

9 An anonymous reviewer has pointed out that the role of working memory in comprehension might be due to a task-related effect. This cannot be excluded. For example, the number of images to consider might originate differences among studies. Engel de Abreu *et al.* (2011) used a task very similar to ours based on a four-picture choice task and they also found a correlation between working memory and comprehension. Montgomery *et al.* 2008, used instead a smaller set of 3 pictures, and no memory effect was detected.

## References

- Adani, F. (2011). Rethinking the acquisition of relative clauses in Italian: towards a grammatically based account. *Journal of Child Language*, 38(1), 141–165.
- Archibald, L. M., and Gathercole, S.E. (2006). Short-term and working memory in specific language impairment. *The International Journal of Language & Communication Disorders*, 41(6), 675–693. <https://doi.org/10.1080/13682820500442602>.
- Arosio, F., Adani, F., and Guasti, M. T. (2009). Grammatical Features in the Comprehension of Italian Relative Clauses by Children. In J. M. Brucart, A. Gavarró, and J. Solà (Eds.), *Merging features: Computation, interpretation and acquisition* (pp. 138–155). Oxford University Press.
- Arosio, F., and Giustolisi, B. (2019). Agreement and interference in direct object clitic production in Italian monolingual children. *Frontiers in Communication*, 3, 66. <https://doi.org/10.3389/fcomm.2018.00066>
- Arosio, F, Guasti, MT, and Stucchi, N. (2011). Disambiguating information and memory resources in children's processing of Italian relative clauses. *Journal of Psycholinguistic Research*, 40(2), 137–54. <https://doi.org/10.1007/s10936-010-9160-0>.
- Baauw, S. (1999). The role of clitic-full pronoun distinction in the acquisition of pronominal reference. Proceedings of the 23rd Boston University Conference on Language Development [BUCDL 23] (pp. 32–43). Somerville, MA: Cascadilla Press.
- Baauw, S., Escobar, L., and Philip, W. (1997). A Delay of Principle B Effect in Spanish Speaking Children: The Role of Lexical Feature Acquisition. In A. Sorace, C. Heycock, and R. Shillcock (Eds.), *Proceedings of the GALA '97 Conference on Language Acquisition*. Edinburgh, Scotland: Human Communication Research Centre.
- Babyonyshev, M., and Brun, D. (2003). The role of aspect in the acquisition of passive constructions in Russian. Paper presented at GALA, Utrecht.
- Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36(3), 189–208.
- Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford Psychology Series.
- Baddeley, A. D., Gathercole, S. E., and Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158–173.
- Baddeley, A. D., and Hitch, G. J. (1974). Working Memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (vol. 8, pp. 47–90). New York: Academic Press
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. In arXiv [stat.CO]. arXiv. <http://arxiv.org/abs/1406.5823>
- Belletti, A. (1999). Italian/Romance clitics: Structure and derivation. In Henk van Riemsdijk (ed.) *Clitics in the languages of Europe* (pp. 543–579). Berlin-New York: Mouton.
- Biondo, N., Pagliarini, E., Moscati, V., Rizzi, L., and Belletti, A. (2022). Features matter: the role of number and gender features during the online processing of subject- and object- relative clauses in Italian. *Language, Cognition and Neuroscience*.
- Bishop, D. V. (2003). *Test for Reception of Grammar (TROG-2)* (2nd ed.). Pearson Education.
- Borer, H., and Wexler, K. (1987). The Maturation of Syntax. In T. Roeper and E. Williams (Eds.), *Parameter setting and language acquisition* (pp. 123–172). Dordrecht: Reidel.
- Boyle, W., Lindell, A. K., and Kidd, E. (2013). Investigating the role of verbal working memory in young children's sentence comprehension. *Language Learning*, 63: 211–242. <https://doi.org/10.1111/lang.12003>

- Brehm, L., and Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, 125, 104334. <https://doi.org/10.1016/j.jml.2022.104334>
- Collins, C. (2005). A smuggling approach to the passive in English. *Syntax*, 8, 81–120. <https://doi.org/10.1111/j.1467-9612.2005.00076.x>
- Delage, H., and Frauenfelder, U.H. (2020). Relationship between working memory and complex syntax in children with Developmental Language Disorder. *Journal of Child Language*, 47(3), 600–632. <https://doi.org/10.1017/S0305000919000722>.
- Dispaldro, M., Ruggiero, A., and Scali, F. (2014) Real-time comprehension of gender and number in four- to seven-year-old children: a study of the relationship between Italian clitic pronouns and visual picture referents. *Journal of Child Language*, 42(4), 917–31. <https://doi.org/10.1017/S0305000914000427>
- Engel de Abreu, P. M. J., Gathercole, S. E., and Martin, R. (2011). Disentangling the relationship between working memory and language: The roles of short-term storage and cognitive control. *Learning and Individual Differences*, 21, 569–574.
- Friedmann, N., Belletti, A., and Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua. International Review of General Linguistics. Revue Internationale de Linguistique Generale*, 119(1), 67–88.
- Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., and Martin, A. J. (1999). Phonological short term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, 13, 65–77.
- Gavarró, A., Torrens, V., and Wexler, K. (2010) Object clitic omission: two language types, *Language Acquisition*, 17(4), 192–219. <https://doi.org/10.1080/10489223.2010.509262>
- Gibson, E. (2000). The Dependency Locality Theory: A Distance-based Theory of Linguistic Complexity. In A. Marantz, Y. Miyashita, and W. O’Neil (Eds.), *Image, language, brain* (pp. 95–126). MIT Press.
- Gordon, P. C., Hendrick, R., and Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1411–1423.
- Grodner, D., and Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science*, 29(2), 261–290.
- Grüter, T. (2006). Object clitics and null objects in the acquisition of French. Ph.D. dissertation, McGill University, Montreal.
- Hamann, C, Kowalski, O., and Philip, W. (1997). The French ‘Delay of Principle B Effect’. In E. Hughes, M. Hughes, and A. Greenhill (Eds.), *Proceedings of the Annual Boston University Conference on Language Development* (p. 21). Somerville, MA: Cascadilla Press.
- Hamann, C., Rizzi, L., and Frauenfelder, U. H. (1996). On the Acquisition of Subject and Object Clitics in French. In H. Clahsen (Ed.), *Generative perspectives on language acquisition* (pp. 309–334). Amsterdam: John Benjamins.
- Harrell, and Harrell. (2019). Package “hmisc” CRAN2018. <ftp://sourceforge.mirror.ac.za/cran/web/packages/Hmisc/Hmisc.pdf>
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, 59(4), 434–446. <https://doi.org/10.1016/j.jml.2007.11.007>
- Jaeggli, O. A. (1986). Passive. *Linguistic Inquiry* 17.4, 587–622.
- Jakubowicz, C. (2005). The Language Faculty: (ab)normal development and interface constraints. Paper presented at GALA 2005, Siena.
- Jakubowicz, C. (2011). Measuring derivational complexity: new evidence from typically-developing and SLI learners of L1-French. *Lingua*, 121, 339–51.
- Kayne, R. (1975). *French Syntax*. Cambridge, Mass.: MIT Press.
- Kim, S. (2015). ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, 22(6), 665–674. <https://doi.org/10.5351/CSAM.2015.22.6.665>
- King, J., and Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30(5), 580–602.
- Leclercq, A. L., and Majerus, S. (2010). Serial-order short-term memory predicts vocabulary development: Evidence from a longitudinal study. *Developmental Psychology*, 46(2), 417–427. <https://doi.org/10.1037/a0018540>

- Lecocq, P.** (1996). *L'ECOSSE: une épreuve de compréhension syntaxo-sémantique*. Lille: Presses Universitaires de Septentrion.
- Lewis, R. L., and Vasishth, S.** (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, **29**(3), 375–419.
- Manetti, C.** (2013). On the Production of Passives in Italian: Evidence from an Elicited Production Task and a Syntactic Priming Study with Preschool Children. In S. Baiz, N. Goldman and R. Hawkes (Eds.), *Proceedings of the 37th annual Boston University Conference on Language Development [BUCLD37]*. Online proceedings supplement. Somerville, MA: Cascadilla Press. <http://www.bu.edu/buclid/supplementvol37/>
- Maracó Coelho de Barros Pereira, R.** (2009). The passive in 3- and 4-year-olds. *Journal of Psycholinguistic Research*, **38**, 435–446.
- Marini, A., Gentili, C., Molteni, M., and Fabbro, F.** (2014). Differential verbal working memory effects on linguistic production in children with Specific Language Impairment. *Research in Developmental Disabilities*, **35**, 3534–3542.
- Marini, A., Marotta, L., Bulgheroni, S., and Fabbro, F.** (2015). *Batteria per la Valutazione del Linguaggio in Bambini dai 4 ai 12 anni (BVL\_4-12)*. Firenze: Giunti O.S.
- Marini, A., Piccolo, B., Taverna, L., Berginc, M., and Ozbič, M.** (2020). The complex relation between executive functions and language in preschoolers with developmental language disorders. *International Journal of Environmental Research and Public Health*, **17**, 1772.
- McKee, C.** (1992). A comparison of pronouns and anaphors in Italian and English acquisition. *Language Acquisition* **1**, 21–55.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D.** The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 2000, **41**, 49–100.
- Miyake, A., and Friedman, N.P.** The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 2012, **21**, 8–14.
- Montgomery, J. W., Magimairaj, B., and O'Malley, M.** (2008). The role of working memory in typically developing children's complex sentence comprehension. *Journal of Psycholinguistic Research*, **37**, 331–354.
- Moscati, V., and Rizzi, L.** (2014). Agreement configurations in language development: a movement-based complexity metric. *Lingua*, **140**, 67–82.
- Moscati V., Rizzi, L., Vottari, I., Chilosi, A., Salvadorini, R., and Guasti, M.** (2020). Morphosyntactic weaknesses in Developmental Language Disorder: The role of structure and agreement configurations. *Journal of Child Language*, **47**(5), 909–944. <https://doi.org/10.1017/S0305000919000709>
- Moscati, V., and Tedeschi, R.**, 2009. The delay of the Italian Past Participle agreement. In Proceedings of the XXXIII Boston University Conference on Language Development (vol. 2, pp. 379–390). Cascadilla Press.
- Pérez-Leroux, A. T., Pirvulescu, M., and Roberge, Y.** (2008). A syntactic transitivity approach to null objects in child language. *Lingua*, **118**(3), 370–398.
- Pierce, A.** (1992). The acquisition of passives in Spanish and the question of A-chain maturation. *Language Acquisition*, **2**, 55–81.
- Prévost, P.** (2009). *The Acquisition of French: The Development of Inflectional Morphology and Syntax in L1 Acquisition, Bilingualism and L2 Acquisition*. Amsterdam: John Benjamins.
- R Core Team.** (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Riva, V., Cantiani, C., Dionne, G., Marini, A., Mascheretti, S., Molteni, M., and Marino, C.** (2017). Working memory mediates the effects of gestational age at birth on expressive language development in children. *Neuropsychology*, **31**(5), 475–485.
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., and Gabrieli, J. D. E.** (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, **29**(5), 700–710.
- RStudio Team.** (2020). *RStudio: Integrated Development for R*. Boston, MA: RStudio, PBC. <http://www.rstudio.com/>.
- Rubin, Maracó Coelho de Barros Pereira.** (2009). The passive in 3- and 4-year-olds. *Journal of Psycholinguistic Research*, **38**, 435–446.

- Schad, D. J., Vasishth, S., Hohenstein, S., and Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, **110**, 104038. <https://doi.org/10.1016/j.jml.2019.104038>
- Schaeffer, J. (2000). *The Acquisition of Direct Object Scrambling and Clitic Placement*. Amsterdam: John Benjamins.
- Serratrice, L., Sorace, A., and Paoli, S. (2004). Crosslinguistic influence at the syntax-pragmatics interface: Subjects and objects in English-Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition*, **7**(3), 183–205.
- Sportiche, D. (1996/98). Clitic Constructions. In J. Rooryck and L. Zaring (Eds.), *Phrase structure and the lexicon* (pp. 213–276). Dordrecht: Kluwer.
- Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, **116**(1), 71–86. <https://doi.org/10.1016/j.cognition.2010.04.002>
- Staub, A., Dillon, B., and Clifton, C. (2017). The matrix verb as a source of comprehension difficulty in object relative sentences. *Cognitive Science*, **41**, 1353–1376. <https://doi.org/10.1111/cogs.12448>
- Tedeschi, R. (2009). *Acquisition at the Interfaces – A case study on object clitics in early Italian* (LOT Dissertation Series), Utrecht.
- Terzi, A., and Wexler, K. (2002). A-Chains and S-Homophones in Children's Grammar: Evidence from Greek Passives. In M. Hirotani (Ed.), *Proceedings of NELS 32*. University of Massachusetts.
- Traxler, M. J., Morris, R. K., and Seely, R. E. (2002). Processing subject and object relative clauses: evidence from eye movements. *Journal of Memory and Language*, **47**(1), 69–90. <https://doi.org/10.1006/jmla.2001.2836>
- van der Lely, H. K. J. (1996). Specifically language impaired and normally developing children: verbal passive vs adjectival passive sentence interpretation. *Lingua*, **98**, 243–72.
- Varlokosta, S., Belletti, A., Costa, J., Friedmann, N., Gavarró, A., Grohmann, K. K., Guasti, M. T., Tuller, L., Lobo, M., Anđelković, D., Argemí, N., Avram, L., Berends, S., Brunetto, V., Delage, H., Ezeizabarrena, M.-J., Fattal, I., Haman, E., van Hout, A., de López, K. J., Katsos, N., Kologranic, L., Krstić, N., Kraljević, J. K., Mićević, A., Nerantzini, M., Queraltó, C., Radic, Z., Ruiz, S., Sauerland, U., Sevcenco, A., Smoczyńska, M., Theodorou, E., van der Lely, H., Veenstra, A., Weston, J., Yachini, M., and Yatsushiro, K. (2016). A cross-linguistic study of the acquisition of clitic and pronoun production. *Language Acquisition*, **23**(1), 1–26. <https://doi.org/10.1080/10489223.2015.1028628>
- Volpato, F., Verin, L., and Cardinaletti, A. (2016). The comprehension and production of verbal passives by Italian preschool-age children. *Applied Psycholinguistics*, **37**(4), 901–931. <https://doi.org/10.1017/S0142716415000302>
- Wechsler, D. (1993). *WISC—R. Scala d'intelligenza Wechsler per bambini riveduta*. Firenze, Italy: Giunti O.S.
- Wexler, Ken. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, **106**, 23–79.
- Zesiger, P., Zesiger, L. C., Arabatzi, M., Baranzini, L., Cronel-Ohayon, S., Franck, J., Frauenfelder, U. H., Hamann, C., and Rizzi, L. (2010). The acquisition of pronouns by French children: A parallel study of production and comprehension. *Applied Psycholinguistics*, **31**(4), 571–603. <https://doi.org/10.1017/S0142716410000147>

## Appendix A. List of the items extrapolated by the Grammatical Comprehension task to assess age-related effects on the comprehension of SVO, passive, and clitic sentences

Item's number	Sentence type	Sentence heard by the child
(8)	SVO	il gatto rincorre il cane <i>"the cat chases the dog"</i>
(23)	SVO	il cane rincorre il gatto <i>"the dog chases the cat"</i>
(25)	SVO	la mamma lava il bambino <i>"the mother washes the boy"</i>
(33)	SVO	il bambino spinge la bambina <i>"the boy pushes the girl"</i>
(35)	SVO	il bambino lava la mamma <i>"the boy washes the mother"</i>
(9)	Passive	la bambina è spinta dal bambino <i>"the girl is pushed by the boy"</i>
(12)	Passive	il bambino è spinto dalla bambina <i>"the boy is pushed by the girl"</i>
(21)	Passive	il gatto è rincorso dal cane <i>"the cat is chased by the dog"</i>
(28)	Passive	il cane è rincorso dal gatto <i>"the dog is chased by the cat"</i>
(15)	3ps Object Clitic	il bambino li tiene <i>"the boy holds them"</i>
(16)	3ps Object Clitic	il bambino la spinge <i>"the boy pushes it"</i>
(17)	3ps Object Clitic	la bambina lo tiene <i>"the girl holds it"</i>
(20)	3ps Object Clitic	la bambina le tiene <i>"the girl holds them"</i>

## Appendix B. Output of the stepwise model selection analysis for memory effects, through the *anova* function.

### Memory forward

m0: Accuracy  $\sim 1 + (\text{Complex\_Structures} + \text{Movement}) * \text{Age} + (\text{Complex\_Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

mf1: Accuracy  $\sim 1 + (\text{Complex\_Structures} + \text{Movement}) * \text{Age} + \text{Forward} + (\text{Complex\_Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

mf2: Accuracy  $\sim 1 + (\text{Complex\_Structures} + \text{Movement}) * \text{Age} + \text{Forward} + \text{Forward} : (\text{Complex\_Structures} + \text{Movement}) + (\text{Complex\_Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m0	13	5184.5	5279.8	-2579.2	5158.5			
mf1	14	5166.6	5269.2	-2569.3	5138.6	19.94	1	<b>7.99e-06</b>
mf2	16	5169.2	5286.6	-2568.6	5137.2	1.33	2	0.52

### Memory backward

m0: Accuracy  $\sim 1 + (\text{Complex\_Structures} + \text{Movement}) * \text{Age} + (\text{Complex\_Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

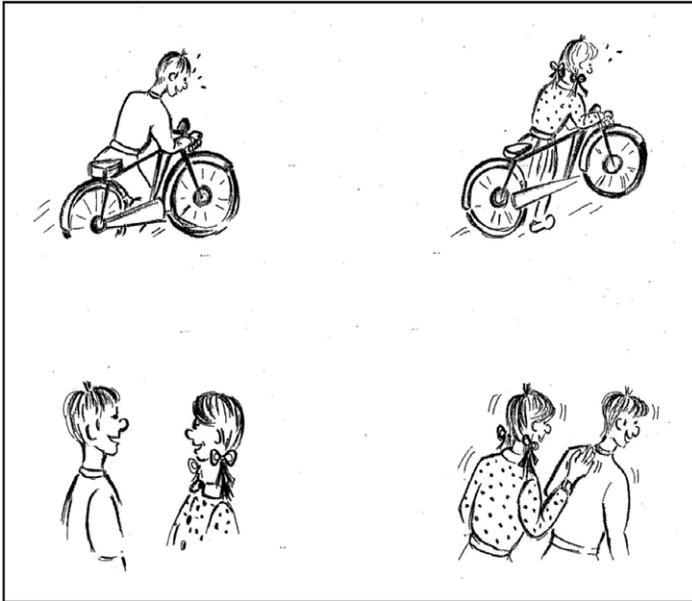
mb1: Accuracy  $\sim 1 + (\text{Complex\_Structures} + \text{Movement}) * \text{Age} + \text{Backward} + (\text{Complex\_Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

mb2: Accuracy  $\sim 1 + (\text{Complex\_Structures} + \text{Movement}) * \text{Age} + \text{Backward} + \text{Backward} : (\text{Complex\_Structures} + \text{Movement}) + (\text{Complex\_Structures} + \text{Movement} | \text{Subj}) + (1 | \text{Item})$

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
mf0	13	5184.5	5279.8	-2579.2	5158.5			
mf1	14	5158.4	5261	-2565.2	5130.4	28.11	1	<b>1.144e-07</b>
mf2	16	5161.6	5278.9	-2564.8	5129.6	0.79	2	0.68

## Appendix C

Item 16: il bambino la spinge  
the boy cl<sub>fem, plur</sub> pushes  
“the boy pushes it” (please note that  
in Italian bicycle is a fem. noun)



**Cite this article:** Moscati, V., Marini, A., & Biondo, N. (2023). What a thousand children tell us about grammatical complexity and working memory: A cross-sectional analysis on the comprehension of clitics and passives in Italian. *Applied Psycholinguistics* 44, 1161–1184. <https://doi.org/10.1017/S0142716423000462>