Complex grammar in English: A snapshot of comprehension in children aged 5 to 8

(Vikki Janke & Gloria Chamorro)

(accepted version to appear in First Language)

Abstract

Our study makes an empirical contribution to questions relating to the developmental trajectory of four examples of English complex grammar: subject and object control, subject and object relative clauses, long passives, and *seem*-raising constructions with and without an overt experiencer argument. We tested children's comprehension of all seven sentence sets at the same point in time using a picture-selection task. 45 children (20 girls) from three Year groups (1, 2, and 3) with a mean age of 6.3, 7.4, and 8.3 years participated. The three groups scored at ceiling on subject relatives and on raising without an experiencer, and there were Year differences in order of age for object and subject control. Subject control showed a predictably delayed pattern and success with it correlated positively with verb-knowledge scores. However, all Year groups performed less well – with no differences between Years – on passives, object relatives, and raising with an experiencer, suggesting that even at age 8, these constructions were not fully comprehended. The most problematic construction was raising with an experiencer, where all Years achieved a mean score of 3/6 or below. We discuss this data pattern in relation to four grammatical properties (empty categories, displacement, intervention, word order), frequency and the lexical idiosyncrasies of some of the verbs. With respect to the grammatical properties, we ask whether certain combinations are more difficult for children to navigate than others.

Keywords

English complex grammar; control; generative grammar; passives; raising; relative clauses

1. Introduction

Very young children's early morphological and grammatical development of English has been tracked quite successfully in that a broadly agreed typical trajectory has been proposed; this includes their progression from one-word to two-word utterances, a relatively predictable order of grammatical morpheme production, and the formation of early question and negative structures (Brown, 1973; Clark, 2016). But from the age of four, the grammatical picture becomes less clear. A broad range of more complex constructions are being negotiated, and it has proven more difficult to pinpoint the order in which these constructions are acquired and to identify the grammatical factors that impact most upon their developmental sequence (see Guasti, 2017)[1]. Several of the complex[2] sentences that children are navigating share key grammatical properties, for example, having a phonetically null argument (called an empty category (*ec*) here) in an embedded clause that needs to be interpreted, as in (1a and b). However, they can differ importantly in other respects, such as whether their derivation includes so-called 'displacement' of that *ec* from the underlying position in which it is interpreted to a different surface one in which it is pronounced, as in (2a and b). A further complicating factor is whether displacement of the *ec* results in a change of order between the subject and object with respect to the transitive verb's thematic relations, as in (3a and b). If this order is disrupted, as in (3b), the word order is described as non-canonical (see explanations in Section 1.1).

(1) a. John told Mary ***ec*** to read the book.

b. John seemed to Mary ***ec*** to be reading the book.

(2) a. John told Mary [***ec*** to read the book]. No *ec* Displacement

[1] This paper is written from a generative grammar perspective and relies on concepts and terminology from this tradition. We have tried to keep technical terms to a minimum and to enhance accessibility, have provided brief explanations and examples as concepts are introduced. An interested reader would find Guasti (2017) a useful place to start for more detailed explanations.

[2] We use the term 'complex' here in a theory-neutral way to refer to sentences that consist minimally of an independent clause and a dependent clause/phrase.

b. John seemed to Mary [***ec*** to be reading the book]. Displacement of *ec*

(3) a. It was John who ***ec*** kissed Mary in the park. Canonical Order

b. It was John who Mary kissed ***ec*** in the park. Non-canonical Order

The present study builds on previous acquisition work undertaken on examples of grammar that are well-known for being acquired later. In the present context, later-acquired refers to sentences that can continue to cause difficulty for typically developing children from five to six years of age and beyond. There is ongoing debate as to these sentences' correct classifications, the grammatical properties they have, and how these impact on the order of acquisition. The answers matter because categorising these sentences with greater precision enables us to edge closer to a clearer formulation of grammatical complexity (see De Cat & Melia, 2022). Equally, with an accurate developmental trajectory of advanced grammar, one can notice signals of delay or disruption in vulnerable children. But completing a picture of grammatical development is exciting, too, because it contributes to linguistic theory, namely the ongoing question of what our underlying linguistic and cognitive apparatus must be like for us to comprehend and produce such sentences. In this article, we have a more modest aim, namely that of gaining a clearer empirical picture of five to eight-year-old children's comprehension of four examples of complex grammar in English: obligatory control, passives, relative clauses, and raising constructions. The study is unique in its probing of children's comprehension of all four construction types at the same point in time using the same task and materials. This is important because direct comparisons between the acquisition paths of different constructions become much harder when those studies employ distinct designs and test materials. Differential performance on different tasks is a well-known phenomenon in the acquisition literature (see Frizelle et al., 2019).

A discussion of the vast acquisition literature on these sentence types is beyond the scope of this article (see Guasti, 2017, for a review). However, on the assumption that sentences with

shared syntactic properties should group similarly in terms of children's performance on a task using the same design and administered under the same conditions, our empirical contribution will feed usefully into ongoing debates over these constructions' trajectory, and relatedly, their underlying characteristics. In the next subsection, we illustrate our test sentences and explain their grammatical properties. Having described some key similarities and differences, we turn in Section 1.2 to the current study and its research questions. Our methods are set out in Section 2, and Section 3 presents our data. Section 4 discusses some key results in light of the properties introduced in Section 1.1.

1.1 The sentence sets

The sentence sets included in this study are double-complement subject- and object-control, as in (4a and b), long passives (so-called because the sentence includes the external argument of the passivized verb in an optional 'by phrase'), illustrated in (5), subject and object relative clauses, shown in (6a and b), and *seem*-raising constructions with and without an overt experiencer argument as per (7a and b).

(4a) **SC**: Luna$_i$ promised Harry *ec*$_i$ to make the cake really chocolatey.

(4b) **OC**: Hermione persuaded Ron$_i$ *ec*$_i$ to read the book of spells.

(5) **P**: Harry$_i$ was watched *ec*$_i$ by Hermione in the kitchen.

(6a) **SR**: It was Harry$_i$ who *ec*$_i$ read to Luna on the green sofa.

(6b) **OR**: It was Ron$_i$ who Hermione kissed *ec*$_i$ by the pond.

(7a) **RNE**: Luna$_i$ seems *ec*$_i$ to be lifting Ron with a magic spell.

(7b) **RWE**: Luna$_i$ seems to Ron *ec*$_i$ to be feeling seasick in the boat.

A brief look at these sentences will highlight four key grammatical properties that link or distinguish them. In SC and OC there is a silent subject i.e. an **EMPTY CATEGORY** in the infinitival subordinate clause, whose reference is determined by a syntactic antecedent; this

antecedent may be the main-clause subject, as in (4a) or the main-clause object, as in (4b). The indices demonstrate this interpretative dependency. Note also that in (4a) the distance between the antecedent and the *ec* to which it is interpretatively linked is further than in (4b), where the object and *ec* are local to one another. That is, in double-complement SC but not OC, an argument **INTERVENES** between the *ec* and its antecedent. Both SC and OC are examples of obligatory control because the referential relation between the antecedent and *ec* is obligatory: the *ec* cannot refer to a different argument in the sentence or to a sentence-external one from the discourse, unlike non-obligatory control, which is pragmatically regulated (see Landau, 2013). With respect to (5), in passive constructions (P), the verb's thematic object is pronounced in the subject position. Aside from this **DISPLACEMENT** between where the argument is interpreted and where it is pronounced, passives also result in a **NON-CANONICAL** word order because the position in which the arguments are pronounced does not correspond to their thematic mapping (i.e. the thematic object of the verb is not in the verb's complement position, as it is in its active counterpart, *Hermione watched Harry in the kitchen*). In these long passives, the thematic subject of the verb *watch* is included optionally in a 'by-phrase'. Example (6) illustrates two subtypes of relative clauses. SR clauses, as in (6a), involve displacement of the subject from the embedded clause to the main clause, but unlike the passive, this displacement has not led to a non-canonical order: the subject still precedes the verb *read* and the object still follows it. There is also no intervening argument between the word *who* and its extraction site. However, in OR clauses, as in (6b), it is the object that is displaced. Thus, OR sentences involve not only *ec*s and displacement but also non-canonical order (the thematic object of *kiss* precedes its subject), and intervention of another argument (the subject intervenes between where the object is interpreted and where it is pronounced). Lastly, in *seem*-raising constructions, as in (7), the thematic subject of the embedded non-finite clause (i.e. *Luna*) is not pronounced in this position but in the subject position of the main clause, which means these constructions also incorporate an *ec* and displacement but not a non-canonical order; these so-called 'subject-to-subject' raising sentences can occur with **no** overt experiencer (RNE) intervening between the *ec* and its

antecedent, as in (7a), or **with** an overt experiencer (RWE) intervening between them, as in (7b). For example, (7b) shows that the argument, *Ron*, intervenes between the *ec* and its antecedent, *Luna*. Table 1 summarises each of these constructions with respect to the properties just described.

Table 1. Construction type and core grammatical properties.

	Empty Category	Displacement	Intervention	Non-Canonical Order
Subject Control	Yes	No	Yes	No
Object Control	Yes	No	No	No
Passives	Yes	Yes	No	Yes
Subject Relatives	Yes	Yes	No	No
Object Relatives	Yes	Yes	Yes	Yes
Raising No Experiencer	Yes	Yes	No	No
Raising With Experiencer	Yes	Yes	Yes	No

Each of the syntactic properties in Table 1 could contribute to the acquisition challenge of a construction and so impact on its developmental trajectory, and an aim of the present study is to observe which constructions appear harder to comprehend. The sentence sets chosen provide us with an opportunity of noting whether comprehension difficulties pattern with certain combinations of the grammatical properties in Table 1. They also lend themselves well to being incorporated into one task, namely picture selection.[3] Below we refer to some key works that have considered these constructions' development, although space considerations and the large number of sentence sets examined in this study mean we cannot do justice to the rich acquisition literature that exists on each. Given this study's focus on English and comprehension, we prioritise works that link most closely to these dimensions.

Research on OC and SC has consistently shown that acquisition of OC precedes SC by quite some years (see Eisenberg & Cairns, 1994; Mateu, 2016). Children as young as four comprehend and produce sentences such as (1a), whereas their acquisition of SC, mostly

[3] Question structures were not included for this reason.

demonstrated with the verb *promise*, exhibits a delay such that children at age seven and eight have still been found to struggle with them (Chomsky, 1969; Janke & Perovic, 2015; Sherman & Lust, 1993). Aside from their difference in terms of locality, where in OC, the interpretive relation between antecedent and *ec* is not interrupted by another argument, OC is also a relatively frequent construction, most often tested with verbs such as *tell*, which appears sooner in young children's vocabulary (Bloom et al., 1984). Both of these are factors that might contribute further to children's earlier grasp of it. In contrast, in SC, the verb *promise* is a so-called 'mental-state' verb, which introduces another potential layer of difficulty to this construction's acquisition, independently of the intervention between antecedent and *ec* set out above (see Nixon, 2005; Pascual et al., 2008). In addition, these kinds of SC constructions are far less frequent. Other verbs that can be used in this sentence frame are limited (e.g. *swear to*, *threaten to, vow to*) and are less likely to form part of a young child's vocabulary so most work on this control-type has been with the verb *promise*. Focusing on English and Spanish, Mateu (2016) compared the pattern of delayed acquisition of SC *promise* with that of *seem*-raising, asking whether it is the presence of an intervener that is at the crux of these constructions' delayed acquisition or whether divergent syntactic derivations for each of these constructions are implied by the children's performance on them, too. If intervention were the sole source of children's delay, we might expect their performance on SC and RWE to be similar, whereas a distinct pattern of performance would at least be consonant with their having diverging derivations. Using a Truth-Value Judgement Task with children aged four to six, results on *promise* in both languages were comparable, where all children fared poorly (with a mean of 46-48% correct) on the relevant SC constructions. However, results on RWE were mixed, where the mean percentage correct was 56% for English-speaking children and 75% for Spanish-speaking children. Interestingly, when children's individual performances on SC and RWE were compared, no correlations were found. That is, children who succeeded with RWE struggled with SC and vice versa, suggesting the difficulties associated with these constructions might stem from different sources.

Work on P sentences in English has shown that children up to the age of five have great difficulty producing and comprehending these (Crawford, 2012; Hirsch & Wexler, 2006). One important distinction that has been drawn is between Ps incorporating actional verbs (e.g. *kiss, push*) and those built around non-actional verbs (e.g. *remember, love*), where correct performance on the former precedes that of the latter across different tasks (Agostinho et al., 2024; Maratsos et al., 1985; Nguyen & Pearl 2021; Oliva & Wexler, 2018). An oft-cited proposal with respect to this distinction is that children might interpret verbal short Ps with actional verbs as adjectival passives, enabling them to bypass the passive derivation (Borer & Wexler, 1987). These adjectival analyses are less easy to employ with non-actional or perception (e.g. *see, watch, hear)* verbs and Ps that include 'by phrases', implying that the more valid test for measuring the acquisition of Ps is one that uses these verbs in long Ps (Perovic & Wexler, 2010). When non-actional and/or perception verbs are used, the age at which these constructions are comprehended is much later, with studies reporting success at ages of six or seven (Baldie, 1976; Hirsch & Wexler, 2006; Perovic et al., 2014). Agostinho et al. (2024) compared Portuguese children's comprehension of long Ps of actional (*comb, paint*) and perception verbs (*see, hear*) using a sentence-picture matching task. They found the predicted contrast between verb types, with children from the age of four gaining just over 80% correct for actional long Ps whereas this percentage was not reached for perception-verb long Ps until eight. Results on *hear*, however, were markedly worse than for *see*, a result that patterns with O'Brien et al. (2006). Looking further at non-actional verbs, they sourced the problem with non-actional verbs to the semantic property of 'affectedness', namely the degree to which the internal argument of the relevant verb is affected by that verb: the more affected the argument, the less difficulty the child should have with it. In support of their predictions, they found that children fared more poorly with *look for* than with *push,* the latter of which has a more affected internal argument, given the physical contact involved. The actional/non-actional distinction has also been looked at more closely by Aravind and Koring (2023), who link the absence of a developmental parallel between these sentence types to Ps that have subject-experiencer predicates. Specifically, their claim is that children misanalyse these

structures as unaccusatives, and since unaccusatives do not passivise (e.g. *The tree was fallen), children do not comprehend them. *See* and *watch* are both perception verbs with similar imageability yet only *see* has an experiencer subject (see footnote 5). It would be interesting, therefore, to see if any contrast were to emerge in the face of these verbs' passive instantiations.

Research on comprehension in SR and OR shows consistently that OR poses greater difficulty not only for children but adults, too. Crosslinguistically, children have been shown to acquire SR first (Adani, 2011; Friedmann et al., 2009; Lau & Tanaka, 2021; Martins et al., 2018), especially in so-called head-first languages, where the verb precedes its complement in the main clause (Tanaka et al, 2024), and adults show faster reading times for SR than OR (Crain et al., 2001). In both SR and OR, the subordinate clause is embedded within the DP that it modifies but in OR, the linear (see Gibson, 2000) and structural (see Collins 1994; Friedmann et al., 2009; Rizzi, 1990) distance between the relativised variable and its antecedent is greater. OR interpretation can be improved by reducing the processing cost associated with the intervening referent, for example, by making the intervenor a pronoun rather than a proper noun (Warren & Gibson, 2002) or by introducing a number mismatch between the intervening argument and the relativised one. This has led to revisions of how early children can comprehend (and produce) ORs (Adani et al., 2010; Diessel & Tomasello, 2005) but this facilitation does not answer for the consistent asymmetry found between SR and OR in comprehension when these factors are held constant (but see Tanaka et al, 2024 for a meta-analysis that compares children's SR and OR preferences in head-first and head-last languages and in languages with head-initial and head-final RCs). Given that OR includes an *ec*, displacement, intervention, and non-canonical order, a difference in performance on these constructions might be expected. What is not clear is if it is the greater number of properties that compounds the learning task, or whether the problem can be reduced to a particular property (see Arosio et al., 2017; Friedmann et al., 2009). Martins et al. (2018) conducted a cross-sectional study using an act-out and sentence-judgement task with Portuguese children

from three to eleven years, comparing performance on RCs with that of obligatory control. Children in the highest age band (8 to 11 years) scored at ceiling on SR, whereas results for OR were much poorer with no significant differences across age groups on either task. For the sentence-judgement task, even the highest age bands' means were not far above chance. Their results on SC were also of note, suggesting a steep learning curve for SC, where only those in the highest age band had a full grasp of SC. This contrasted with the much flatter incline in improvement for OR, pointing to different developmental paths - a possibility that could be explored with a longitudinal design. These results tie in with a theoretical analysis of control that does not involve displacement (e.g. Williams, 1980; Landau, 2000; Janke, 2007), implying that the operation of crossing an intervening argument through displacement, as in OR, creates a persistent challenge for children, and that this is different to the exceptional 'breaking of locality' that children encounter with SC-*promise* (see Janke & Perovic, 2015).

As opposed to a large consensus over the order of development of OR and SR in head-initial languages such as English, the literature on raising is less clear. Like SC, raising with the verb *seem* is widely reported as a more difficult structure (Choe & Deen, 2016; Orfitelli, 2012), with some studies showing children still struggling at age nine (Hirsch, 2011; Koring, 2007). However, different accounts make different predictions about children's performances on RWE versus RNE. Whereas Hirsch (2011) and Mateu (2016) report children at age six struggling similarly with RWE and RNE, Becker (2006), Choe (2012), and Choe and Deen (2016) have reported that children acquire RNE sooner. Some of these results have been discussed in relation to processing costs, and also whether RNE sentences do indeed contain no experiencer in their syntactic representation, or whether they in fact contain a syntactically present but phonetically null experiencer, i.e. a further *ec*. Theories that postulate a syntactic *ec* expect children to perform similarly on (7a) and (7b), whereas those that do not, predict children to find (7b) more difficult than (7a) (see Orfitelli, 2012; Snyder & Hyams, 2015). In the aforementioned study by Mateu (2016), English-speaking children's performance on RNE (*The dog seems to be grey*) and RWE (*The dog seems to the cat to be grey*) remained

problematic for six-year-olds, whereas Spanish-speakers comprehended RNE (*El perro parece ser gris)* at age four. Mateu's interpretation was that in English, both *seem*-raising constructions have a syntactic intervener, even if covert in the case of RNE, whereas in Spanish, RNE lacks an intervener altogether. The mixed results reported with respect to RNE and RWE motivate further examination of children's performance on these constructions.

This brief review of these seven sentence sets has shown that they share the fact that they have clauses with at least one *ec*. After this, they start to differ. Some involve displacement and others, on most accounts (cf Hornstein, 1999), do not (e.g. raising versus control). We have seen constructions that involve displacement which result in a non-canonical word order and others where displacement does not (e.g. OR and P versus SR). Still other constructions have arguments that intervene between the *ec* and its antecedent as a consequence of displacement, which contrast with constructions where intervention occurs without displacement (e.g. RWE and OR versus SC). Thus, when examined together using the same materials, children's performance on this collection of sentence sets provides a unique opportunity of observing which (combination of) syntactic obstacles prove more difficult than others. In the next section, we set out the present study, which examines five-to-eight-year-old English-speaking children's comprehension of them.

1.3 Aims of the study

Our study examined comprehension and production of seven sentence sets (OC, SC, OR, SR, P, RNE, and RWE) at the same point in time, with the same children, by administering a picture-selection task and a sentence-repetition task. Here we focus on the picture-selection. We also administered a verb-knowledge task, which was relevant to the OC constructions as some mental-state verbs were used (i.e. *promise* and *persuade*), and lexical knowledge of these verbs could have impacted performance on the OC constructions independently of the grammatical properties distinguishing the constructions (see Janke & Perovic, 2015; Sherman

& Lust, 1993). Our aim was to take a developmental snapshot of all seven constructions. The main research questions were:

(I) How did children's comprehension performance on the seven sentence sets (OC, SC, OR, SR, P, RNE, RWE) differ across the three Year groups?

(II) For object control, how did children's performance differ with respect to the critical verbs (i.e. *persuade/tell/order*)? In particular, was performance with *tell* better than with the mental-state verb *persuade*?

(III) For passives, how did children's performance differ with respect to the critical verbs (*see/listen/watch*)? In particular, was performance with *watch* better than with the experiencer predicate, *see*?

(IV) To what extent did children's performance on SC and OC correlate positively with their understanding of *promise* and *persuade* respectively?

2. Method

2.1 Participants

45 children attending state-funded mainstream schools from the Southeast region of England, spread equally across three school Years took part, ranging from 5.7 to 8.9 years. None had special educational needs or hearing impairments, and all were described by their teachers and parents as neurotypical. Table 2 illustrates the gender distribution, mean age and age range per Year group.

Table 2. Number of participants, mean age, and age range per Year group.

	N (girls)	Mean age (SD)	Age range
Year 1	15 (6)	6.31 (0.32)	5.71-6.78
Year 2	15 (7)	7.43 (0.27)	6.99-7.79
Year 3	15 (7)	8.31 (0.22)	8.03- 8.90

2.2 Procedure

Ethical approval was obtained from the University of Kent, and informed consent was obtained from the schools and parents. Children also gave their assent verbally before each session. They undertook the tasks individually in a quiet room at school. Five tasks were administered in two different sessions[4], with each session lasting no longer than 20 minutes and taking place one week apart. During the first session, children completed the picture-selection task and the British Picture Vocabulary Scales (BPVS). In the second, they undertook a sentence-repetition task, a working-memory (WM) task, and a verb-knowledge task. Children were given a sticker and a small gift for taking part.

2.3 Materials

Picture-selection task

We used an untimed, two-choice, picture-selection task, adapted from Janke & Perovic (2017a, 2017b). Viewed as a low-demand task which facilitates a larger number of items per condition, variations of it have been used successfully with children with and without clinical diagnoses (Janke & Perovic, 2015; Sanoudaki & Varlokosta, 2014), and it is also associated with less between-subject variation (see Adani, 2011). For each trial, children saw two pictures and chose the one that best matched the accompanying sentence; this appeared visually at the bottom of the screen and auditorily through headphones. The sentences were prerecorded in a soundproof booth, using a female native English speaker, who maintained a nuclear stress. Item presentation was randomised automatically for each child, and location of the correct picture was balanced throughout (left or right), as were the figures in the pictures. To reduce task demands and complications further, we avoided definite DPs (i.e. the boy) and inanimate referents, and used proper nouns based on four key characters from the Harry Potter series (Harry, Ron, Hermione, and Luna). All test items were semantically reversible in

[4] We included a measure of receptive vocabulary (BPVS-3, Dunn et al., 2009) and an auditory forward and backward digit span task (WISC-IV, Wechsler, 2004) in our test battery. These related to the sentence-repetition task so are not discussed further here.

that either character in each sentence could stand for the Agent/Experiencer/Theme/Patient depicted, and for each item, the two arguments were represented by one female and one male character. The task included seven critical sentence types (OC, SC, OR, SR, P, RNE, RWE), with six tokens of each. The test sentences were between 8 and 11 words long and included adverbials at the ends of sentences (e.g. in the pond, last night, carefully) to help even the length[5]. Two control sentence sets were incorporated (one using *seem* in a non-raising context and one using *think* set against pictorial depictions that had been used in the RWE/RNE and P sentences). Ten fillers were also included, giving a total of 64 trials. A trial was scored as 1 if the correct picture was selected or 0 if the wrong picture was selected. Ten adults piloted the task, all of whom scored 64/64.

Sentence types used for picture-selection task

The complete list of sentences used in the task are available on the Open Science Framework repository (https://osf.io/ts6f4/); here, we provide one example of each construction tested. Where possible, our choice of verbs was guided by their imageability. For OC, (8), the main-clause verbs were *persuade*, *order,* and *tell*, and the verbs in the complement clause were *read*, *fly,* and *row*. Each instantiation appeared twice with the main characters alternating between agent and patient roles. For SC, (9), the verb in the main clause was *promise* and the verbs in the complement clause were *pop*, *mak*e, and *drink*. Both types of RC sentences used cleft sentences with an expletive followed by a copula linking to the focused predicate nominal preceding the relative clause (e.g. It was Peter who played the piano), which are more likely to occur in young children's early RC productions. RCs with centre-embeddings were also avoided (see Diessel & Tomasello, 2005; Kidd et al., 2007). OR sentences, (10), included the verbs *kiss*, *splash,* and *tap*, and the SR sentences, (11), used the verbs *lift*, *read,* and *tap.* The three verbs chosen to test P, as per (12), were *see, watch,* and *listen*. All three are verbs

[5] The mean number of words of the critical sentences for each condition was 8.7 (SD=0.52) for OC, 9.6 (SD=1.03) for SC, 9 (SD=0.89) for OR, 9.6 (SD=1.37) for SR, 9 (SD=0.89) for P, 8.7 (SD=1.03) for RNE, and 10.3 (SD=0.51) for RWE.

of perception but whereas *see* is a stative verb whose thematic subject is an experiencer, *watch* and *listen* are not stative and their subjects are thematic agents[6]. Two types of raising were tested: RNE, as in (13), and RWE, as per (14). Both employed *seem* as the critical raising verb and the verbs in the infinitival clauses were *lift*, *make*, *play*, *splash*, *feel,* and *pass.*

(8) Harry ordered Luna to fly the broomstick properly. (OC)

(9) Harry promised Hermione to make the cake really chocolatey. (SC)

(10) It was Harry who Luna splashed in the potions class. (OR)

(11) It was Ron who read to Hermione on the green sofa. (SR)

(12) Luna was seen by Harry outside the school gates. (P)

(13) Luna seems to be lifting Ron with a magic spell. (RNE)

(14) Hermione seems to Harry to be making a delicious cake. (RWE)

To depict *seem* in the raising examples, we used images denoting feelings to avoid using thought bubbles (i.e. ascending hearts to illustrate appreciation of 'a lovely tune' and a child rubbing their stomach and drooling to depict appreciation of 'a delicious cake'), and we monitored understanding of these depictions via our control sets. The degree to which the meaning of *seem* has been confused by children has also been discussed as poor understanding of the concept would affect task performance for reasons orthogonal to the raising operation. Thus, we included a set of control items, such as (15), that tested *seem* in unraised contexts, and a set that included the verb *think*, which checked children's knowledge of thought/mind, which is integral to their ability to interpret the meaning of the word *seem* (see Hirsch, 2011).

(15) It seems that Ron is kissing Hermione on the cheek. (Control Seem)

[6] *Watch* and *listen* pattern differently from *see* on tests for agency *(Luna *saw/watched/listened to Harry deliberately and *See her!/watch her!/listen to her!)* but none of these perception verbs have resultative states *(e.g. Luna saw/watched/listened to Harry for an hour versus *Luna saw/watched/listened to Harry in an hour),* i.e. they are atelic.

(16) Hermione thinks that the cake looks delicious. (Control Think)

Ten fillers, which used a range of constructions, including perception verbs used in the critical P sentence sets, were also incorporated; examples are given in (17), (18) and (19).

(17) Luna is listening to Ron in the music room.

(18) Harry is feeding the owl but Hermione is not watching him.

(19) Luna is feeling seasick while Ron is rowing the boat.

Verb-knowledge task

Because the control constructions involved mental-state verbs, we assessed knowledge of *persuade* and *promise*, using a short task adapted from Janke and Perovic (2015), which asked children two questions about these verbs' meanings. The first was an open question (what does it mean when you persuade someone/promise someone something?), and the second, a bolstered one, used an example (if you persuade your friend that s/he should eat the cake, does that mean that s/he eats it or that s/he does not eat it?/if you promise your mum that you will tidy up your room, does that mean that you do it or you do not do it?). Each child received a score between 0 and 2 for each question (i.e. 0-4 for each verb). The task with sample answers and coding can be accessed via the following link: https://osf.io/ts6f4/.

3. Results

3.1 Control and filler sentences

Performance on the control sentences and fillers were analysed first. All children performed at ceiling (see Table 3).

Table 3. Means, SDs, and Medians of control *seem*, control *think*, and fillers by Year

	Control *seem* (max score = 6)	Control *think* (max score = 6)	Fillers (max score = 10)

	M (SD)	Mdn	M (SD)	Mdn	M (SD)	Mdn
Year 1	5.73 (0.46)	6	5.80 (0.41)	6	9.93 (0.26)	10
Year 2	5.80 (0.41)	6	5.93 (0.26)	6	9.80 (0.41)	10
Year 3	5.87 (0.35)	6	6 (0)	6	9.87 (0.35)	10

We then ran a generalised linear model (GLM) with a logistic link function to see if task order affected the results but the means for Orders 1 and 2 were identical, both having 74% correct answers (*Wald Chi-Square*(1)=.014, *p*=.907).

3.2 The critical sentences

To address Research Question 1, we explored how performance on each of the experimental constructions differed between Years. One-way ANOVAs showed a significant main effect of Year for OC (*F*(2,42)=3.42, *p*=.042). Comparisons across Years showed that Year 3 outperformed Year 1 (*p*=.014). A marginal main effect of Year was found for SC (*F*(2,42)=2.98, *p*=.06), with Year 3 outperforming Year 1 (*p*=.02). No other constructions indicated significant differences between the three Year groups. Each Year's mean scores are illustrated in Table 4, and Figure 1 plots these means for each construction by Year.

Table 4. Mean scores (SDs) for the experimental constructions by Year (maximum score = 6 for each set of constructions).

	RNE	SR	OC	SC	P	OR	RWE
Year 1	5.93 (0.26)	5.93 (0.26)	5.33 (0.82)	3.40 (2.26)	3.27 (1.71)	3.60 (1.64)	2.53 (1.92)
Year 2	5.80 (.56)	5.73 (0.80)	5.73 (0.70)	4 (1.56)	4.33 (1.95)	3.20 (1.26)	3 (1.81)
Year 3	5.87 (0.52)	5.87 (0.35)	5.93 (0.26)	5.07 (1.79)	4 (2)	4.33 (1.63)	2.93 (2.40)

[insert Figure 1]

Figure 1. Mean scores for the experimental constructions by Year.

As Table 4 and Figure 1 indicate, all Years scored most highly on RNE, SR, and OC, with mean scores well above 5 out of 6 for all three constructions. RWE was the most difficult construction, with all three Years gaining mean scores of 3 (out of 6) or below. The order of performance for some of the constructions is different in each Year. Table 5 presents the descending order of performance on each construction according to Year.

Table 5. Descending order of performance on each construction by Year (1=highest score; 7=lowest score).

	1	2	3	4	5	6	7
Year 1	RNE/SR	RNE/SR	OC	OR	SC	P	RWE
Year 2	RNE	SR/OC	SR/OC	P	SC	OR	RWE
Year 3	OC	RNE/SR	RNE/SR	SC	OR	P	RWE

Group performance on each construction was further explored with a series of paired t-tests (Table 7 in the Appendix lists the t-test results for each Year). For Year 1, scores for RNE and SR, which were identical, were significantly higher than scores for all other structures, their OC scores were significantly higher than OR, SC, P, and RWE, and their OR scores were significantly higher than SC, P, and RWE. For Year 2, RNE scores were significantly higher than P, SC, OR, and RWE, their OC and SR scores, which were identical, were significantly higher than P, SC, OR, and RWE, and their P scores were significantly higher than RWE. For Year 3, OC scores were significantly higher than OR, P, and RWE, their SR and RNE scores, which were identical, were significantly higher than OR, P, and RWE, and their SC and OR scores were significantly higher than RWE. Overall, the picture that emerges for all Years is that OC and SR were significantly easier than OR, and RNE was significantly easier than RWE. A similar significant advantage was found for OC and SR over SC in Years 1 and 2. For Year 3, these differences were not significant as the children scored very highly on all three constructions, although numerically OC (5.93) and SR (5.87) were still higher than SC (5.07).

Our analyses for Research Questions 2 and 3 considered how certain verbs affected children's performance on the OC and P constructions. A series of pairwise comparisons were run and, for OC, which incorporated the verbs *persuade*, *order,* and *tell*, *persuade* resulted in significantly better performance (98% correct) than *order* (91% correct) (p=.049). But performance with *tell* (95% correct) was not significantly different from *persuade* (p=.246) or *order* (p=.386). For P, which used *watch*, *listen,* and *see*, the verb *watch* resulted in significantly better performance (72% correct) than *listen* (58% correct) (p=.039). However, performance with *see* (62% correct) was not significantly different from *watch* (p=.148) or *listen* (p=.539).

Lastly, with respect to Research Question 4, we wanted to examine the relation between scores with *persuade* and *promise* on the verb-knowledge task and those same verbs on the picture-selection task. Turning to the verb-knowledge task first, Table 6 shows that the children in Year 1 have the lowest means and that they are comparable for both verbs. In Year 2, the mean for persuade is lower than it is for promise. The children in Year 3 have the highest scores, and the means for the two verbs are identical.

Table 6. Mean scores (SDs) on the verb-knowledge task for *persuade* and *promise* by Year group (maximum scores for each = 4).

	Persuade	Promise
Year 1	2.13 (0.64)	2.8 (1.01)
Year 2	2.6 (0.74)	3.27 (0.96)
Year 3	3.47 (0.64)	3.47 (1.18)

The relation between knowledge of *persuade* and performance with this verb in the picture selection task could not be calculated due to ceiling effects with it in the latter task. For *promise*, however, there was a significant positive correlation between performance with this

verb on the picture-selection and verb-knowledge task (*Kendall's tau correlation*=.348, *p*=.006).

4. Discussion

Our study looked at comprehension of complex grammar in English in typically-developing children across three Years (1, 2 and 3, with mean ages of 6.3, 7.4 and 8.3 respectively), using a two-choice picture-selection task. Four main construction types were tested, three of which included two subtypes, resulting in seven sentence sets: Object Control (OC), Subject Control (SC), Object Relatives (OR) Subject Relatives (SR). The children scored at ceiling on RNE and SR, and near ceiling on OC. The construction that caused the most difficulty for all children, with all Years achieving a mean score of 3 out of 6 or below, was RWE. OR, P, and SC also caused difficulties for children across Years. However, with SC, there was a significant improvement with each successive Year that was not found for OR, P, or RWE. We discuss these findings in turn, starting with the more expected findings first.

The OC results are entirely in line with the literature on this construction's acquisition. With a mean age of 6.3 years, we expected even the youngest Year to perform well with this construction, their being slightly older than the age at which children can still struggle with OC (Eisenberg & Cairns, 1994; Lust, 1986). But despite their high scores, this Year group did score significantly worse than the oldest one, which, with a mean age of 8.3 years, performed at ceiling level. Because of the children's very high scores on OC, a correlation between the children's knowledge of the verb *persuade* and their success with this mental-state verb on the picture-selection task could not be calculated. However, comparisons between the children's results for each verb showed that they were more accurate with *persuade* (98%) than with *order* (91%) and *tell* (95%), although the difference was only significant between *persuade* and *order*. What this shows is that the widely reported earlier onset of OC, based

largely on studies using verbs such as *tell,* which have a high frequency and early acquisition onset, is maintained when using less frequent and mental-state verbs, such as *order* and *persuade* respectively*,* even with children just under six (see also Janke, 2018). In contrast to their competence with OC, the children illustrated a predictably delayed pattern with SC. There was a marginal difference between Years, with Year 1 scoring numerically lower than Year 2, who in turn, scored lower than Year 3, but it was only between Years 1 and 3 that these differences were significant; by Year 3, children gained a group mean score of 5 out of 6 (see Figure 1), which although very high is still not at ceiling. The positive correlation between children's success with *promise* and their scores on the verb-knowledge task, points to understanding of this mental-state verb being part of the explanation for this construction's notoriously late acquisition (see Lust, 1986; Martins et al., 2018). Viewed in conjunction with the infrequency of this construction and the fact that children must (exceptionally) break locality to master it - the matrix object intervenes between the *ec* and the subject antecedent – its more staggered trajectory is understandable. Note, however, that on most theoretical accounts, control does not incorporate displacement of the *ec* whereas raising does (see Landau, 2013, for a review), which makes the children's results on SC particularly interesting in comparison with their results on RWE (Janke & Perovic, 2015; Mateu, 2016). Unlike SC, performance on RWE did not yet show any improvement with age. All Year groups performed equally poorly on this construction, with none surpassing a mean score above 3 out of 6, which suggests a developmental asynchrony between these constructions, a possibility that a longitudinal study could confirm. Recall that although intervention is a property of both RWE and SC, on most accounts, displacement only occurs in RWE (see examples 4a and 7b), making it possible that the children's particularly poor results in RWE reflect this extra dimension. However, similar to *promise*, *seem* is also a mental-state verb, so children's grasp of this verb independently of the purported raising operation is important to consider, too. Unlike *promise*, *seem* cannot be a noun so children have fewer opportunities to explore its meaning, and it is also a verb that lacks a thematic subject, making it more difficult to incorporate into a verb-knowledge task effectively. However, *seem* can be used in unraised

sentences, so it was important that the children completed, and gained ceiling scores on, the six unraised control items. In addition, the children's ceiling scores on the *think* control items supports their having the concept underlying *seem*. At this point, their results on RNE also become pertinent: with no (overt) intervening argument, the children performed at ceiling on this construction, suggesting that it is the presence of displacement in addition to intervention that exacerbates the challenge in *seem*-raising constructions. In this respect, our results are in line with Becker (2006) and Choe (2012), who also reported more difficulties with RWE than with RNE. The possibility of RNE constructions containing a covert intervenor has also been discussed (see Orfitelli, 2012), however, the sharp distinction in the current children's performance on the two sentence sets speaks more strongly in favour of an English RNE representation in which there is no covert syntactic argument.

Interestingly, displacement in the absence of intervention is also what characterises SR, another set on which children scored at ceiling, so they patterned together on two constructions (SR and RNE) that share the presence of two properties (*ec*, displacement) and the absence of another (intervention). In fact, there is a further property these sentences lack, namely a non-canonical word order, which brings us to the children's relative clause results. The OR sentences were the only set which include all the grammatical properties noted in Table 1. If difficulty increases purely as a function of the number of properties, one might have expected children to fare particularly poorly on this construction. Of first note is the stark difference between their SR and OR results, where SR was at ceiling but OR was one of the constructions all groups scored poorly on. These data align with the vast literature on these constructions' comprehension, which has reported that the former are acquired sooner than the latter (Adani, 2011; Martins et al., 2018; Lau & Tanaka, 2021) so our experimental paradigm has corroborated another widely reported pattern. There were no Year group differences for OR, indicating that even at 8 years (when clues, such as contrasting the relevant arguments in terms of animacy or number are not given), English-speaking children can still struggle with ORs even when they are not centre-embedded. Note that the OR pattern

is also different from SC, where children's scores indicted near full comprehension at age 8, and improved in line with their developing lexicon, as indicated by their results on the verb-knowledge task. For OR, however, the problem seems to be a protracted one, a result also reported in Martins et al. (2018). OR, then, with its constellation of grammatical properties is predictably hard but not more so than P or RWE, which have fewer of the syntactic obstacles in Table 1 than OR.

Let us revisit the results on the P sentences. There were no Year differences for P. In fact, the mean Year scores for P and OR were the same, ranging from 3.2 to 4.3 out of 6 (as per Table 5), indicating a similar level of difficulty for both sentence sets despite displacement in P targeting an argument position and displacement in OR targeting a non-argument one. We used long passive sentences with perception verbs, which circumvented children potentially achieving correct answers without computing the passive derivation. Both actional verbs and the optional 'by phrases' have been argued to make this a possibility (Armon-Lotem et al., 2016; Gordon & Chafetz, 1990; Perovic et al., 2010; Snyder & Hyams, 2015). Of the three verbs we used, *watch* resulted in significantly higher accuracy than *listen* (72% versus 58%), the latter of which is conceptually similar to the verb *hear* and has a record of poor performance in tasks testing the passive (see Agostinho et al., 2024; Maratsos et al., 1985; O'Brien et al., 2006). However, the verb *see* (with 62%) did not differ significantly from either verb. Recall also that *see* differs from *watch* in that the former's thematic subject is an experiencer whereas the latter's subject is an agent (see footnote 6). Given the similarity between these two verbs in imageability and the fact that for none of the perception verbs were the internal arguments depicted as being affected (see Agostinho et al., 2024; Ambridge et al., 2016), one might have expected, as per Aravind and Koring (2023), a better performance with *watch.* However, although numerically in the expected direction, there were no significant differences between

them. Furthermore, *listen*, whose subject also behaves as an agent, achieved the lowest score[7].

It could be that the preposition, *to*, accompanying *listen* introduced a further layer of difficulty but verbs with prepositions (i.e. *read to*) were also used in SR, which achieved ceiling level scores. *Listen to* (with the same picture from the P trial) was also used in one of the fillers (*Luna is listening to Ron in the music room*), and the same picture occurred in one of the 'control *think*' trials (*Hermione thinks that the song is beautiful*) and ceiling scores were obtained for both. The lower scores for '*listen to'* will have contributed to children's overall performance on P but they were not significantly different from *see* so the P results cannot be fully explained by that particular verb. Returning to the syntactic properties in Table 1, P sentences do include an *ec*, displacement, and non-canonical word order but they do not incorporate intervention, while RWE, on which all Years struggled most, has an *ec,* displacement, and intervention but not non-canonical word order. These results could point to a derivation comprising displacement and intervention (as in RWE) being trickier for children than one comprising displacement and non-canonical word order (as in P). However, this raises the question of why children fared worse on RWE than they did on OR, which has all the properties under discussion. Frequency is likely to form part of the answer. RWE appears in far fewer contexts than OR, which has numerous guises. RWE, for example, is possible with the verb *appear* but most other examples of raising are RNE, namely verbs that do not select an experiencer, as with the verbs *tend, or happen* or raising adjectives such as *likely to*, *certain to*, giving the language-learning child far fewer opportunities to experience the RWE sentence frame. Mateu (2016), in her study on raising in English and Spanish, compared the frequency of adults' and children's productions of RWE and RNE in the CHILDES corpus

[7] In our test materials, it could be that the passivised argument would be described as more affected in the *listen* and *watch* trials than in the *see* trials since there is a possibility in the *see* trials that the character depicting that argument could be unaware of having been seen. For *watch*, a character is being watched while baking in the kitchen and for *listen*, a character is being listened to while playing the piano. For *see*, the relevant character is seen flying their broomstick.

(MacWhinney, 2000), and reported that when adults used *seem* in a raising context, it was in an RNE frame for 96% of productions and in an RWE frame for only 4%. This huge difference was mirrored by the children for whom the relevant percentages were 94.1% and 5.8% respectively (Mateu, 2016, p. 70). In contrast, verbs that occur with OR are not restricted to a few exemplars, nor do they have the special thematic properties of *seem.* There are also several OR subtypes to which children can be exposed:

(20) It was the boy who/that the girl chased.

(21) The boy who/that the girl chased ran across the playground.

(22) The teacher watched the boy that/who the girl chased.

(23) The boy the girl chased is my friend.

These extra hurdles attached to RWE, which are absent from OR, go some way to accounting for the particularly low scores of RWE relative to OR and P, enabling us to return to the proposition that displacement coupled with non-canonical order presents the language-learning child with the harder challenge.

5. Summary

Our study took a snapshot of the comprehension of seven constructions in the same children between the ages of five and eight. For those constructions where the comprehension literature is broadly in agreement over their developmental trajectory, as with OC versus SC, and OR versus SR, our data aligned with previous results. But where discussions continue over children's order of development and over the sentences' underlying grammatical properties, such as with SC versus RWE, and RWE versus RNE, our data make an empirical contribution to current questions. Firstly, the potential developmental asynchrony between SC and RWE, for example, feeds into discussions over the syntactic derivations underlying obligatory control and raising. The results here are in line with an analysis of control that does not include displacement. Secondly, the contrast in results between RWE and RNE is relevant

to continuing debates over whether there is a syntactically present yet covert intervenor in the latter construction. The current results suggest that in English there might not be. Lastly, the relatively poor performance across the board on P and OR in comparison to SR points to displacement and non-canonical order being the two common denominators that, in combination, present the most difficult learning challenge even if the displacement targets different syntactic positions. It will be interesting to see how these children's results on comprehension of these sentence sets map with their production of them, an investigation we leave for a future study.

Acknowledgements

Foremost, our thanks go to the schools, parents and children without whom this project could not have gone forward. For help with recruitment, we're particularly grateful to Alison Blackwell at Graveney Primary school and to Rina Kreitman. Thanks also to Liane Chester, who assisted with task building, to Gordon Craig for statistical advice, to the British Academy for funding the project and lastly to three anonymous reviewers for critical and constructive comments.

Funding

This research was funded by a British Academy Grant (SRG2223\230847) awarded to both authors.

References

Adams, A.-M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders, 35*(1), 95–116. https://doi.org/10.1080/136828200247278

Adani, F. (2011). Re-thinking the acquisition of relative clauses in Italian: Towards a grammatically-based account. *Journal of Child Language, 38*(1), 141–165. http://dx.doi.org/10.1017/S0305000909990250

Adani, F., van der Lely, H. K.J., Forgiarini, M., & Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua, 120*(9), 2148–2166. http://dx.doi.org/10.1016/j.lingua.2010.03.018

Agostinho, C., Santos, A. L., & Duarte, I. (2018). The acquisition of control in European Portuguese. In A. L. Santos, & A. Gonçalves (Eds.), *Complement clauses in Portuguese: Syntax and acquisition* (pp. 263–294). John Benjamins Publishing Company.

Agostinho, C., Gavarró, A.., & Santos, A. (2024) The acquisition of the verbal passive: The role of verb type. *Language Acquisition*, 1-22. https://doi.org/10.1080/10489223.2024.2331238

Ambridge, B., Bidgood, A., Pine, J., Rowland, C., & Freudenthal, D. (2016). Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies. *Cognitive Science,* 40, 1435–1459. https://doi.org/10.1111/cogs.12277

Aravind, A., & Koring, L. (2023). Experiencer troubles: A reappraisal of the predicate based asymmetry in child passives. *Language Acquisition*, *30*(1), 76–100. https://doi.org/10.1080/10489223.2022.2115373

Armon-Lotem, S., Hamann, E., Jensen de Lopez, K., Smoczynska, M., Yatsushiro, K., Szczerbinski, M., van Hout, A., Dabasinskiene, I., Gavarro, A., Hobbs, E., Kamandulytė-Merfeldienė, L., Katsos, N., Kunnari, S., Nitsiou-Michaelidou, C., Sundahl Olsen, L., Parramon, X., Sauerland, U., Torn Leesik , R., & van der Lely, H. (2016). A large scale cross-linguistic investigation of the acquisition of passive. *Language Acquisition, 23*(1), 27–56. https://doi.org/10.1080/10489223.2015.1047095

Arosio, F., Panzeri, F., Molteni, B., Magazù, S., & Guasti, M. (2017). The comprehension of Italian relative clauses in poor readers and in children with SLI. *Glossa: a journal of general linguistics, 2*(1), 9. https://doi.org/10.5334/gjgl.107

Baldie, B. J. (1976). The acquisition of the passive voice. *Journal of Child Language, 3*(3), 331–348. https://doi.org/10.1017/S0305000900007224

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*(5–6), 507–584. https://doi.org/10.1080/016909697386628

Becker, M. (2006). There began to be a learnability puzzle. *Linguistic Inquiry, 37*(3), 441–456.

Bloom, L., Tackeff, J., & Lahey, M. (1984). Learning to in complement constructions. *Journal of Child Language*, 11, 391—406.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper, & E. Williams (Eds.), *Parameter Setting* (pp. 123–172). Springer.

Brown, R. (1973). *A First Language: the early stages*. Allen and Unwin.

Choe, J. (2012). *Children seem to know raising: Raising and intervention in child language*. University of Hawai'i at Mānoa PhD dissertation.

Choe, J., & Deen, K. (2016). Children's difficulty with raising: A performance account. *Language Acquisition, 23*(2), 112–141. https://doi.org/10.1080/10489223.2015.1047097

Chomsky, C. (1969). *The Acquisition of Syntax in Children from 5 to 10*. MIT Press.

Clark, E. (2016). *First Language Acquisition* (3rd). Cambridge University Press.

Crain, S., Ni, W., & Shankweiler, D. (2001). Grammatism. *Brain and Language, 77*(3)**,** 294–304. https://doi.org/10.1006/brln.2000.2403

Crawford, J. (2012). *Developmental perspectives on the acquisition of the passive.* University of Connecticut PhD dissertation.

De Cat, C., & Melia, T. (2022). What does the sentence structure component of the CELF-IV index, in monolinguals and bilinguals? *Journal of Child Language, 49*, 423–450.

Diessel, M., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, *81*(4), 882–906. https://dx.doi.org/10.1353/lan.2005.0169

Dunn, L. M., Dunn, D. M., & Styles, B. (2009). *The British Picture Vocabulary Scale - Third Edition (BPVS 3).* GL Assessment.

Eisenberg, S., & Cairns, H. (1994). The development of infinitives from three to five. *Journal of Child Language, 21*(3), 713–734. https://doi.org/10.1017/S0305000900009521

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order*. The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 234–250. https://doi.org/10.1080/713755604

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*(1), 67–88. https://doi.org/10.1016/j.lingua.2008.09.002

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: a study of SLI and normal development. *Journal of Child Language*, *31*(3), 661–681. https://doi.org/10.1017/s0305000904006269

Frizelle, P., Thompson, P., Duta, M., & Bishop, D. V. M. (2019). Assessing children's understanding of complex syntax: A comparison of two methods. *Language Learning, 69*, 255–291. https://doi.org/10.1111/ lang.12332

Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition, 36*(3), 227–254. https://doi.org/10.1016/0010-0277(90)90058-R

Guasti, M. T. (2017). *Language acquisition: The growth of grammar*. MIT Press.

Hanten, G., & Martin, R. (2000). Contributions of phonological and semantic short-term memory to sentence processing: Evidence from two cases of closed head injury in children. *Journal of Memory and Language, 43*(2), 335–361. https://doi.org/10.1006/jmla.2000.2731

Hirsch, C. (2011). *The acquisition of raising*. MIT PhD dissertation.

Hirsch, C., & Wexler, K. (2006). By the way, children don't know by. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *BUCLD 30: Proceedings of the 30th annual Boston University Conference on Language Development* (pp. 249–261). Cascadilla Press.

Hirsch, C., & Wexler, K. (2007). The late acquisition of raising: What children seem to think about seem. In S. Dubinsky, & W. D. Davies (Eds.), *New horizons in the analysis of control and raising* (pp. 35–70). Springer.

Hornstein, N. (2001). *Move! A minimalist theory of construal*. Blackwell Publishers.

Hyams, N., & Snyder, W. (2005). *Young children never smuggle: Reflexive clitics and the Universal Freezing Hypothesis* [Conference session]. 30th Boston University Conference on Language Development. Boston, MA: United States. http://web.uconn.edu/snyder/papers/HyamsSnyder05.pdf

Isaacs, E. B., & Vargha-Khadem, F. (1989). Differential course of development of spatial and verbal memory span: A normative study. *British Journal of Developmental Psychology*, *7*(4), 377–380. https://doi.org/10.1111/j.2044-835X.1989.tb00814.x

Janke, V. (2007) *Control without PRO*. UCL PhD Dissertation.

Janke, V. (2018) Discourse effects on older children's interpretations of complement control and temporal adjunct control. *Language Acquisition*, 25 (4). pp. 366-391.

Janke, V., & Perovic, A. (2015). Intact grammar in HFA? Evidence from control and binding. *Lingua, 164*, 68–86. https://doi.org/10.1016/j.lingua.2015.06.009

Janke, V., & Perovic, A. (2017) Contrasting complement control, temporal adjunct control and controlled verbal gerund subjects in autism spectrum disorder: The role of contextual cues in reference assignment. *Frontiers in Psychology*, 8: 448.

Janke, V., & Perovic, A. (2017). Advanced syntax and primary pragmatics in children with ASD. In L. Naigles (Ed), *Innovative Investigations of Language in Autism Spectrum Disorder. Language and the Human Life Span Series* (pp. 141–506). American Psychological Association/De Gruyter Mouton, Washington.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: a cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*(6), 860–897. https://doi.org/10.1080/01690960601155284

Koring, L. (2007). *Raising-verbs in Dutch: Structure and acquisition*. Utrecht University MA dissertation.

Lau, E., & Tanaka, N. (2021). The subject advantage in relative clauses: A review. *Glossa: a journal of general linguistics, 6*(1), 1–34. https://doi. org/10.5334/gjgl.1343

Landau, I. (2000). *Elements of Control: Structure and Meaning in Infinitival Constructions.* Springer.

Landau, I. (2013). *Control in Generative Grammar: A Research Companion*. Cambridge University Press.

Lust, B. (1986). *Studies in the Acquisition of Anaphora. Volume I. Defining the constraints.* Kluwer Academic Publishers.

MacWhinney, B. (2000). *The CHILDES project: The database. Vol. 2*. Psychology Press.

Maratsos, M. P., Fox, D. E. C., Becker, J. A., & Chalkley, M. A. (1985). Semantic restrictions on children's passives. *Cognition, 19*(2), 167–192. https://doi.org/10.1016/0010-0277(85)90017-4

Marchman, V.A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, F9-F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

Martins, A., Santos, A. L., & Duarte, I. (2018). Comprehension of relative clauses vs. control structures in SLI and ASD Children, In A. B. Bertolini, & M. J. Kaplan (Eds.), *BUCLD 42: Proceedings of the 42nd annual Boston University Conference on Language Development* (pp. 493–506). Cascadilla Press.

Mateu, V. (2016). *Intervention Effects in the Acquisition of Raising and Control: Evidence from English and Spanish*. University of California PhD dissertation.

McDaniel, D., Cairns, H., & Hsu, J. (1990). Control principles in the grammar of young children *Language Acquisition, 1*(4), 297–335. https://doi.org/10.1207/s15327817la0104_1

Nixon, S. (2005) Mental state verb production and sentential complements in four-year-old children. *First Language, 25*(1), 19–37. https://doi.org/10.1177/0142723705046898

Nguyen, E., & Pearl, L. (2021). The link between lexical semantic features and children's comprehension of English verbal be-passives. *Language Acquisition,* 28(4), 433–450. https://doi.org/10.1080/10489223.2021.1939354

O'Brien, K., Grolla, E., & Lillo-Martin, D. (2006). Long passives are understood by young children. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *BUCLD 30: Proceedings of the*

30th annual Boston University Conference on Language Development (pp. 441–451). Cascadilla Press.

Oliva, J., & Wexler, K. (2018). The acquisition of Spanish passives: The comparison between subject experiencer versus actional verbs and direct aspectual semantic evidence for the adjectival interpretation. In A. Gavarró (Ed.), *On the acquisition of the syntax of romance* (pp. 57–90). John Benjamins Publishing Company.

Orfitelli, R. (2012). *Argument intervention in the acquisition of A-movement*. UCLA PhD dissertation.

Pascual, B., Aguado, G., Sotillo, M., & Masdeu, J. C. (2008). Acquisition of mental state language in Spanish children: a longitudinal study of the relationship between the production of mental verbs and linguistic development. *Developmental Science, 11*(4), 454–466. https://doi.org/10.1111/j.1467-7687.2008.00691.x

Perovic, A., Vuksanovic, J., Petrovic, B., & Avramovic-Ilic, I. (2014). The acquisition of passives in Serbian. *Applied Psycholinguistics, 35*(1), 1–26. https://doi.org/10.1017/S0142716412000240

Perovic, A., & Wexler, K. (2010). Development of verbal passive in Williams syndrome. *Journal of Speech, Language and Hearing Research, 53*(3), 1294–1306. https://doi.org/10.1044/1092-4388(2009/07-0188)

Sanoudaki, E., & Varlokosta, S. (2014). Pronoun comprehension in individuals with Down syndrome: Deviance or delay? *Journal of Speech, Language and Hearing Research, 57*(4), 1442–1452. https://doi.org/10.1044/2014_JSLHR-L-13-0035

Sherman, J. C, & Lust, B. (1993). Children are in control. *Cognition, 46*(1), 1-51. https://doi.org/10.1016/0010-0277(93)90021-M

Snyder, W., & Hyams, N. (2015). Minimality effects in children's passives. In E. Di Domenico, C. Hamann, & S. Matteini (Eds.), *Structures, strategies and beyond: Studies in honor of Adriana Belletti* (pp. 343-368). John Benjamins Publishing Company.

Stadtmiller, E., Lindner, K., Süss, A., & Gagarina, N. (2022). Russian-German five-year-olds: what omissions in sentence repetition tell us about linguistic knowledge, memory skills and their interrelation. *Journal of Child Language*, 49, 869 - 896.

Tanaka, N., Lau, E., & Lee, A. L. F. (2024). On the universality of the subject preference in the acquisition of relative clauses across languages. *First Language*, *44*(3), 219-243. https://doi.org/10.1177/01427237241226734

Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology*, *141*, 65-82. https://doi.org/10.1016/j.jecp.2015.06.015

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children* – Fourth UK Edition (WISC-IV UK). Pearson Clinical.

Williams, E. (1983). Predication. *Linguistic Inquiry* 11, 203–238.

Willis, C. S., & Gathercole, S. E. (2001). Phonological short-term memory contributions to sentence processing in young children. *Memory*, *9*(4–6), 349–363. https://doi.org/10.1080/09658210143000155

Appendix.

Table 7. Paired t-test results for each Year group's performance on the 7 constructions.

	Year 1			Year 2			Year 3	
	t	*df*	*p*	*t*	*df*	*P*	*t*	*df*
RNE – SR	.000	14	1.000	.435	14	.670	.000	14
RNE – OC	2.806	14	**.014**	.323	14	.751	.435	14
RNE – OR	5.534	14	**.000**	7.756	14	**.000**	3.525	14
RNE – SC	4.164	14	**.001**	5.281	14	**.000**	1.666	14
RNE – P	6.162	14	**.000**	2.797	14	**.014**	3.500	14
RNE – RWE	6.859	14	**.000**	5.957	14	**.000**	4.725	14
SR – OC	2.806	14	**.014**	.000	14	1.000	.564	14
SR – OR	5.857	14	**.000**	7.536	14	**.000**	4.075	14
SR – SC	4.528	14	**.000**	4.516	14	**.000**	1.922	14
SR – P	6.325	14	**.000**	3.004	14	**.009**	3.690	14
SR – RWE	6.859	14	**.000**	6.045	14	**.000**	5.047	14
OC – OR	4.133	14	**.001**	8.264	14	**.000**	3.781	14
OC – SC	4.005	14	**.001**	4.377	14	**.001**	1.818	14
OC – P	4.681	14	**.000**	3.146	14	**.007**	4.005	14
OC – RWE	4.836	14	**.000**	5.909	14	**.000**	4.743	14
OR – SC	.379	14	.710	1.517	14	.152	1.434	14
OR – P	1.234	14	.238	1.887	14	.080	.674	14
OR – RWE	1.481	14	.161	.468	14	.647	3.309	14
SC – P	.238	14	.815	.617	14	.547	1.980	14
SC – RWE	1.086	14	.296	1.563	14	.140	3.872	14
P – RWE	.991	14	.338	2.197	14	**.045**	1.697	14