



Sentence repetition task for European Portuguese: results from a study with monolingual and Portuguese-German bilingual children

Liliana Correia, Maria Lobo & Cristina Flores

To cite this article: Liliana Correia, Maria Lobo & Cristina Flores (03 Jun 2024): Sentence repetition task for European Portuguese: results from a study with monolingual and Portuguese-German bilingual children, Language Acquisition, DOI: [10.1080/10489223.2024.2346586](https://doi.org/10.1080/10489223.2024.2346586)

To link to this article: <https://doi.org/10.1080/10489223.2024.2346586>



© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#) 



Published online: 03 Jun 2024.



[Submit your article to this journal](#) 



[View related articles](#) 



[View Crossmark data](#) 

Sentence repetition task for European Portuguese: results from a study with monolingual and Portuguese-German bilingual children

Liliana Correia ^a, Maria Lobo ^b, and Cristina Flores ^a

^aUniversidade do Minho; ^bUniversidade NOVA de Lisboa

ABSTRACT

The Sentence Repetition Task (SRT) is an assessment tool for children's language abilities, which has been used in various languages and speaker types. The present study adapted the LITMUS-SRT to European Portuguese (EP) and applied it to 43 monolingual and 25 bilingual heritage speakers of EP with German as societal language (aged 6 to 10 years), in order to evaluate their knowledge of various syntactic properties, which differ in their level of complexity. Additionally, it also assessed the effect of language experience variables—that is, *richness of the heritage language (HL) input* and *cumulative amount of HL exposure*—on the performance of the bilingual children in the task. Results demonstrate that group (monolingual vs. bilingual), the children's age and the level of complexity (with three levels) predict children's accuracy scores. Furthermore, it is particularly in the highest levels of complexity that the bilingual children show more variation and more protracted development, with clitics and subjunctives being the most challenging linguistic properties. As for the role of the input-related variables, richness of the HL input, but not cumulative amount of exposure, emerged as a significant predictor of the bilingual children's accuracy.

ARTICLE HISTORY

Received 20 October 2022

Accepted 16 April 2024

1. Introduction

It is by now a well-established fact that bilingual and monolingual children go through comparable trajectories of language development and may reach identical levels of proficiency in a target language; however, bilingual children show much more intra-group variation regarding the pace and the outcome of the acquisition process, compared to age-matched monolingual peers (De Houwer 2009). This is due to the variety of extra-linguistic factors that modulate bilingual language development, in particular age of onset of bilingualism, quantity and quality of language input, as well as type and degree of instruction in each language (Hoff et al. 2014, Rodina et al. 2020, Unsworth 2013; see also Paradis 2023, for a recent overview). In addition, language-external variables are in interplay with linguistic factors such as the (dis)similarity between the two (or more) language systems and internal characteristics of each language (Torregrossa et al. 2023). Understanding how language-external and language-internal factors interact in (bilingual) language development is a crucial endeavor for researchers, which adds to our knowledge of the human language faculty.

One important step in this endeavor consists of finding the appropriate instruments to assess children's language abilities, in particular by developing assessment tools that are suitable to test

CONTACT Cristina Flores  cflores@elach.uminho.pt  Centro de Estudos Humanísticos, Universidade do Minho, Escola de Letras Artes e Ciências Humanas, Braga 4710-057, Portugal

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/10489223.2024.2346586>

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

mono- and bilingual children alike (Chiat et al. 2013). This becomes even more challenging when we assess bilinguals' weaker languages, which generally are not the school language and may be characterized by low (or non-existent) literacy skills. This is normally the case of heritage languages (HLs), that is, the family language of bilingual children with an immigration background, who grow up with another dominant environment language. Against this backdrop, the present paper presents an assessment tool for European Portuguese (EP) and the results of its application to a group of monolingual EP children and another group of EP-German bilingual children, who acquired Portuguese as their HL in Germany. We developed a Sentence Repetition Task (SRT) to assess children's knowledge of various syntactic properties, which differ in their level of complexity. We will show that the SRT is suitable to capture the language-internal differences between the tested properties in monolingual language development and, in a more pronounced way, in heritage Portuguese. Given that most work on the syntactic development of (European) Portuguese has focused on monolingual children, this study contributes to this field of research by analyzing the performance of bilingual children as well.

2. Background

2.1. *Role of language exposure in heritage language acquisition*

Research on heritage bilingualism has shown that child HSs' language acquisition environment is more heterogeneous than that of their monolingually-raised peers (Montrul 2016, Paradis 2023). This is because bilingual children's language exposure is divided between the two languages and there is huge individual variation regarding the amount and type of language input they receive from each one (Unsworth 2015).

The variability regarding the language environment in which children acquire the HL has been robustly shown to have an impact on its development across linguistic domains (Flores et al. 2017, Haman et al. 2017, Hoff et al. 2012, Rodina et al. 2020, Torregrossa et al. 2023). In fact, several studies in the last decades have revealed that the pace and outcome of HL acquisition are interconnected with language experience variables such as the amount of current and/or over time HL exposure (Komeili et al. 2020, Mitrofanova et al. 2018, Rodina & Westergaard 2017) as well as the richness of the HL input (Ibrahim et al. 2020, Jia & Paradis 2015). For example, Mitrofanova et al. (2018) investigated Norwegian-Russian bilingual children's acquisition of grammatical gender in HL-Russian, having found that cumulative length of exposure to the HL was the language experience variable that predicted children's performance in grammatical gender tasks the best. Komeili et al. (2020), in turn, found that Farsi-English bilingual children's performance in a SRT in HL-Farsi was significantly and positively correlated with their total use of this language. Ibrahim et al. (2020) also investigated the role of language experience variables in the performance of child HSs of Arabic with German as societal language on a SRT in the HL. They found that length of Arabic-schooling, current Arabic use, and Arabic-linguistic richness were significant predictors of the children's accuracy scores in the task.

Thus, taking the effect of language experience variables into consideration when analysing HL development is an indispensable requirement.

2.2. *Assessing syntactic development with Sentence Repetition Tasks*

Assessing the syntactic development of heritage bilingual children is a challenging task, especially when their HL is the one being assessed, given that they may lack lexical knowledge to complete the task successfully. The SRT is a method that has been successfully used to assess syntactic development in bilingual children, in their majority language (see Almeida et al. 2017, for French in Arabic-French, EP-French and Turkish-French children; Fleckstein et al. 2018, for French in Arabic-French and English-French children; Prentza et al. 2022, for Greek in Albanian-Greek children), as well as in their HL (see Haman et al. 2017, for Polish in Polish-English children; Komeili et al. 2020, for Farsi in Farsi-English children) or, comparatively, in both their languages (see Andreou et al. 2021, for the use of

SRTs in Italian and Greek; Hamann et al. 2020, for German and Arabic; Meir, Walters & Armon-Lotem 2016, for Russian and Hebrew). SRT is an elicited imitation task that does not require as much lexical knowledge as other production tasks and may thus be more appropriate for bilingual children, especially those that are not dominant in the language that is being assessed. Furthermore, the literature convincingly shows that the SRT is a valid method for assessing syntactic knowledge, provided sentence length and other details, such as frequency of lexical items, are controlled for (Blume & Lust 2017, Marinis & Armon-Lotem 2015). Children are not mere parrots that repeat sentences: to repeat a sentence correctly, besides holding the information in memory,¹ the child must process it and comprehend its grammar (Lust et al. 1996). Specific mistakes that children make when they fail to repeat correctly give us information on their implicit grammatical knowledge. It is expected that children will only be able to repeat the sentence correctly when its syntactic structure is part of their internal grammar.

Elicited imitation tasks may be developed to investigate specific linguistic variables in experimental tasks contrasting sentences that differ minimally in specific syntactic aspects (Blume & Lust 2017, Lust et al. 1996). However, elicited imitation tasks have also been used to assess syntactic knowledge of monolingual and bilingual children in clinical settings (Almeida et al. 2017, Marinis & Armon-Lotem 2015), as well as to determine the grammatical knowledge, the language proficiency level and the language dominance of bilingual children or adult second language learners (Andreou et al. 2021, Erlam 2006, Gaillard & Tremblay 2016, Schönström & Hauser 2022, Spada et al. 2015).

For Portuguese, although some language batteries include the elicited imitation of sentences (e.g., SINTACS: Vieira 2011; ALPE: Mendes et al. 2014), there is no SRT based on theoretically informed linguistic criteria which systematically includes linguistic structures of different levels of difficulty designed for school-aged children. Furthermore, to our knowledge, no SRT tool has yet been used to assess bilingual children with EP as HL. Within the COST Action IS0804 ‘Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment,’ versions of the LITMUS-SRT have been developed for more than twenty languages (Marinis & Armon-Lotem 2015).

In this paper, we describe the development and first results of a SRT for EP, adapted from the LITMUS-SRT to assess the grammatical knowledge of monolingual and bilingual children. Its reliability for clinical contexts has not been investigated yet.

In the next section, we describe the different syntactic structures that were included in the EP version of the SRT, considering previous findings on their acquisition, in order to justify their inclusion in our EP version of the task, according to different levels of difficulty.

2.3. Syntactic development in EP

The development of our Sentence Repetition Task-European Portuguese (SRT-EP) followed the rationale for the LITMUS-SRT. It was based on a first version of the task developed within the COST Action IS0804 that was later modified and improved according to criteria that will be explained in Section 3.

The task includes different types of syntactic structures with different levels of difficulty, some of which have been shown to be problematic for children with language impairment or to develop late both in monolingual and bilingual acquisition. Following the rationale of the LITMUS-SRT (Marinis & Armon-Lotem 2015), the SRT-EP includes easier and more complex structures distributed across

¹Sentence repetition is claimed to be a multifaceted task which involves linguistic representations in *long-term memory* (also called *language knowledge*, Ebert 2014) and the maintenance of information in *short-term memory* or *working memory* (Riches 2012). The role that each type of memory plays in SRTs is part of an ongoing debate, with some authors highlighting the association between sentence repetition and linguistic processing and other authors highlighting its relationship with working memory capacity (see Ebert 2014, Okura & Lonsdale 2012, Riches 2012). Nevertheless, despite the varying emphasis given to the role of long-term memory, linguistic processing and working memory capacity, SRTs, when properly constructed (e.g., sentences being long enough to exceed memory capacity and, consequently, disallow passive copying), have been argued to draw on linguistic knowledge (see discussion in Marinis & Armon-Lotem 2015).

three levels of complexity. The complexity of the task is progressive, starting with simpler structures and ending in more complex ones. It includes monoclausal sentences and biclausal sentences with coordination and subordination processes, structures that have proven to be difficult for children in many languages, as well as structures that are challenging in EP.

In this section, we consider previous findings on the monolingual or bilingual acquisition of these structures, that include: simple clauses (with the indicative or the subjunctive); accusative clitics in different syntactic contexts; actional and non-actional passives; finite and infinitival complement clauses; coordinate clauses; relative clauses; *wh*-questions; and conditional structures.

As for monoclausal sentences, the task includes structures with different degrees of complexity: SVO sentences with the indicative, SVO sentences with the subjunctive (triggered by the modal adverb *talvez* ‘maybe’), passives, monoclausal sentences with accusative clitics, and monoclausal sentences with *wh*-movement. In the monoclausal sentences, we expect children to perform better in sentences with the indicative (preterite and periphrastic future) than in sentences with the subjunctive, thus the former are included in complexity level 1 (C1) and the latter in complexity level 2 (C2). This division is based on previous studies, which have reported protracted development of the subjunctive in monolingual and bilingual language acquisition, compared to indicative forms in various syntactic contexts (Flores et al. 2017, 2019; Gonçalves et al. 2011; Jesus et al. 2019). The use of the subjunctive with *talvez* has not been studied yet.

As for the monoclausal sentences with an accusative clitic, these have been shown to be particularly challenging in EP, thus being included in complexity level 3 (C3). EP grammar has accusative clitics that may occur in preverbal position (proclisis) or postverbal position (enclisis) depending on several factors, in particular the presence or absence of a proclisis trigger² (see, for a comprehensive description, Martins 2013) (see [1]):

- (1) a. A mãe do menino lavou-o.
 the mother of.the boy washed-him.CLIT
 ‘The boy’s mother washed him.’
- b. A mãe do menino não o lavou
 the mother of.the boy not him.CLIT washed
 ‘The boy’s mother did not wash him.’

Furthermore, when the referent for the object is linguistically or pragmatically salient, accusative clitics may alternate with a null object in some syntactic contexts, including simple sentences [see (2)]. This object omission construction, however, is not tolerated by most EP native speakers in island contexts [see (3)] (Raposo 1986):

- (2) a. Onde guardaste as chaves? – Guardei na gaveta.
 where kept.2SG the keys? – kept.1SG in.the drawer
 ‘Where did you keep the keys? I kept [them] in the drawer.’
- b. Onde guardaste as chaves? – Guardei-as na gaveta.
 where kept.2SG the keys? – kept.1SG-them.CLIT in.the drawer
 ‘Where did you keep the keys? I kept them in the drawer.’

²Proclisis triggers include negation, negative subjects, some preverbal adverbs, some quantified subjects, complementizers, preposed *wh*-constituents, among others.

- (3) Sabes das chaves?
 ‘Do you know about the keys?’
- a. – *Estou aborrecido porque perdi.
 – am.1SG upset because lost.1SG
 ‘I am upset because I lost them.’
- b. – Estou aborrecido porque as perdi
 – am.1SG upset because them lost.1SG
 ‘I am upset because I lost them.’

These two properties—variable clitic placement and alternation with null objects—make EP clitics particularly difficult to acquire. Although clitics are produced early, their target production in obligatory contexts stabilizes late, as shown in many studies (Costa & Lobo 2007, Costa et al. 2015, Flores et al. 2020, Lira et al. *in press*, Varlokosta et al. 2016). Late mastery of accusative clitics has also been observed for bilinguals and heritage speakers of EP (Costa et al. 2016; Flores & Barbosa 2014; Flores et al. 2016; Nardelli & Lobo 2018; Rinke & Flores 2014; Tomaz et al. 2019, 2020). In EP, children tend to omit accusative clitics at much higher rates than in other languages and until later, and they have trouble dealing with the variable placement pattern of EP: they tend to generalize enclisis to proclitic contexts.

Passive structures have also shown to be difficult for children in many languages (although not all), and they seem to be particularly affected in children with developmental language disorders. So-called “long passives” (the ones with a by-phrase) (4b) and non-actional passives (5) seem to be especially demanding for children:

- (4) a. O cão foi mordido.
 the dog was bitten
 b. O cão foi mordido pelo gato.
 the dog was bitten by.the cat
 ‘The dog was bitten (by the cat).’
- (5) a. O bombeiro foi visto
 the fireman was seen
 b. O bombeiro foi visto pelo rapaz.
 the fireman was seen by.the boy
 ‘The fireman was seen (by the boy).’

These difficulties have been attributed to different factors: non-canonical word order (Bever 1970); semantic properties (Maratsos et al. 1985); development of the mechanism of theta role transmission (Fox & Grodzinsky 1998); the maturation of different grammatical constraints (Borer & Wexler 1987, Hirsch & Wexler 2006, Snyder & Hyams 2015, Wexler 2004).

For Portuguese, different studies have reported late mastery of reversible passives (Sim-Sim 2006), low production of passives in spontaneous production (Estrela 2013), asymmetries between actional and non-actional passives – the latter being more difficult (Agostinho 2020, Estrela 2013), and better comprehension of eventive passives than resultative and stative passives (Agostinho 2020, Corrêa et al. 2017, Estrela 2013, Lima Júnior et al. 2018). Whatever the source that explains their late mastery, passives, and in particular long non-actional passives, are structures that stabilize late and that can be indicators of language development. Based on these previous findings, short passives have been included in C2 and long passives in C3.

The task also includes monoclausal sentences with wh-movement. Wh-questions in EP usually show movement of the wh-constituent to the left periphery (although wh-in situ is also possible). When the moved wh-constituent is the subject, the canonical word order is maintained, whereas when the moved wh-constituent is the object, the canonical word order is disrupted, and the subject intervenes in the chain formed by the moved element and its base position:

- (6) a. Quem é que ~~quem~~ empurrou o polícia?
 who is that ~~who~~ pushed the policeman
 'Who pushed the policeman?'
 b. Quem é que o polícia empurrou ~~quem?~~
 who is that the policeman pushed ~~who~~
 'Who did the policeman push?'
- (7) a. Que homem é que ~~que homem~~ empurrou o polícia?
 which man is that ~~which man~~ pushed the policeman
 'Which man pushed the policeman?'
 b. Que homem é que o polícia empurrou ~~que homem?~~
 which man is that the policeman pushed ~~which man~~
 'Which man did the policeman push?'

There is now ample crosslinguistic evidence that object *wh*-questions are harder for children than subject *wh*-questions, especially when the *wh*-constituent has a lexical restriction, as in (7b) (Bentea 2017, Friedmann & Novogrodsky 2011, among others). EP is no exception: both Cerejeira (2009) and Baião & Lobo (2014) show that EP children produce less target object *wh*-questions than subject *wh*-questions; and both studies also show that children have more trouble comprehending *which* object *wh*-questions than *who* object *wh*-questions.

According to Bentea et al. (2016), following on Friedmann et al. (2009) and Grillo (2009), this stems from a constraint—featural relativized minimality—which causes processing costs for children when there is an inclusion relation between the features of the moved element and the intervener. These asymmetries in the production and comprehension of object *wh*-questions (and, in particular, the ones with a lexical restriction) are quite widespread crosslinguistically, and therefore it is expected that object *which*-questions will cause difficulties for younger/immature children and for monolingual and bilingual children with DLD independently of their dominant language.

Alternative approaches considering adult processing of other structures with *wh*-movement, and relative clauses in particular, have proposed that different factors may play a role, including memory, interpretative processes and frequency (see Gordon & Lowder 2012, for a revision). Specifically, models that consider similarity-based interference during memory retrieval (Gordon et al. 2001, 2004; Lewis et al. 2006; among others), similarly to featural relativized minimality accounts, predict that the processing costs will be higher when the moved element and the intervening element are more similar in their semantic characteristics. In fact, Villata et al. (2016), despite following a featural relativized minimality approach, leave open the possibility of integrating featural relativized minimality accounts and memory-based models. Both approaches predict that subject dependencies will be processed more easily than object dependencies and, crucially, that the processing costs will be modulated by the degree of similarity between the moved element and the intervening element.

Based on these previous studies, we included *who* object *wh*-questions in C2 and *which* object *wh*-questions in C3.

In addition to monoclausal sentences with different degrees of complexity, the task includes biclausal sentences. The production of sentences with two or more clauses, combined through coordination or subordination processes, is usually taken as an indicator of language development (Peristeri et al. 2017).

Biclausal sentences may display different degrees of complexity depending on the type of dependency between the clauses, the depth of embedding of the clause, the presence of *wh*-movement, and semantic properties of the clause. Common coordinators emerge earlier than subordinators (Costa et al. 2008, Costa 2010), and different types of subordinate clauses may display different degrees of

complexity. For instance, relative clauses are embedded sentences that show a higher syntactic complexity than other embedded sentences because they imply wh-movement. These subordinate clauses, and in particular the ones with object dependencies, as argued above for wh-object questions, are difficult for children with language disorders and may be difficult as well for L2 learners and for heritage speakers (Delage & Tuller 2010, Novogrodsky & Friedmann 2006, Scheidnes et al. 2009).

Previous research shows that coordinators emerge earlier than subordinators. Costa et al. (2008) establish the following emergency scale of connectives in EP children's spontaneous productions: *e* 'and' > *mas* 'but' > *porque* 'because' > *se* 'if'. Coordination processes with connectors³ are generally assumed to be easier than subordination processes, and coordinate clauses are much more frequent than subordinate clauses in children's early written productions (Batalha et al. 2022). Therefore, we included coordinate clauses introduced by *e* 'and' and *mas* 'but' in level C2.

For relative clauses, there is also robust crosslinguistic evidence that object relatives (8b) are harder than subject relatives (8a) (Adani et al. 2010, Durrleman & Bentea 2021, Friedmann et al. 2009, among others). Similar results have been obtained for monolingual acquisition of EP. As in other languages,

- (8) a. O gato que mordeu o cão é preto.
 the cat that bit the dog is black
 'The cat that bit the dog is black.'
- b. O gato que o cão mordeu é preto.
 the cat that the dog bit is black
 'The cat that the dog bit is black.'

object relative clauses in EP are harder to produce and to comprehend than subject relative clauses, and this capacity develops over time (Costa et al. 2011, Lobo & Soares-Jesel 2017, Vasconcelos 1993). The source of the asymmetry between the acquisition of subject relatives and object relatives, in addition to non-canonical word order (Bever 1970), has been given different explanations. Some approaches have attributed it to featural relativized minimality, and to the processing cost of structures in which the moved constituent and an intervening constituent share similar features (Friedmann et al. 2009). Alternatively, following studies on adult processing, we can find memory-based models that explain these asymmetries with similarity-based interference or models that consider an interplay between several factors (see Gordon & Lowder 2012, for a review). These different approaches are not mutually exclusive and may in fact be complementary. Whatever is the right explanation, it is a well-grounded fact that object relatives are generally harder than subject relatives. Based on these findings, we included subject relative clauses in C2 and object relative clauses in C3.

As for complement clauses, the task includes both finite and infinitival complement clauses. Finite complement clauses emerge early, according to Santos (2017) and Soares (2006). However, not all properties of finite complement clauses are mastered early. Mood selection (indicative or subjunctive) in the finite embedded clause is subject to development, and children still have problems with some contexts at age 8-9 (Jesus et al. 2019). Protracted development of the subjunctive in embedded complement clauses has also been found in bilingual heritage speakers of EP (Flores et al. 2017, 2019). As for infinitival complement clauses, previous research has found that subject control infinitival clauses selected by volitional predicates are acquired early, whereas other infinitival complement clauses develop later (Agostinho 2014, Santos et al. 2016, Santos 2017) and may remain unstable in adult heritage grammars (Barbosa et al. 2018). For this reason, subject control infinitival clauses selected by volitional predicates were included in C1 and finite complement clauses in C2.

³This may be different for coordination structures without connectors. For instance, juxtaposition contexts that express causal relations have been shown to be more demanding than coordinated sentences with connectors, as well as causal subordinate clauses (Aguilar 2017).

Conditional clauses introduced by *se* ‘if’ emerge later than other types of embedded clauses (Batalha et al. 2022, Costa et al. 2008). Although these embedded sentences do not display movement, they involve a high degree of complexity for syntactic, semantic, and morphological reasons. Some of these adjunct (non-selected) embedded sentences display the subjunctive mood. Additionally, conditional structures can be associated to different semantic values: hypothetical or unreal conditionals (with the subjunctive mood) are harder than real conditionals (with the indicative mood).

- (9) a. Se a criança estiver com febre, leva-a ao médico.
 If the child is.FUT.SUBJ with fever take her to.the doctor
 ‘If (by any chance) the child has fever, take her to the doctor.’
- b. Se a criança estivesse com febre, eu levava-a ao médico.
 If the child was.IMP.SUBJ with fever I would.take her to.the doctor
 ‘If the child had fever, I would take her to the doctor.’
- c. Se a criança está com febre, leva-a ao médico.
 If the child is.IND with fever take her to.the doctor
 ‘If (indeed) the child has fever, take her to the doctor.’

Although there is not much investigation on the acquisition of conditional clauses in Portuguese, the literature reports that different properties of concessive and conditional clauses develop late (Costa 2010, Gonçalves et al. 2011), thus these sentences with conditional predicates were assigned to C3.

Torregrossa et al. (2023) have shown that assessment tools such as cloze tests, which include structures with different degrees of difficulty are useful instruments to capture developmental differences between heritage speakers of EP with different ages, varying degrees of language exposure and in contact with different majority languages. Since cloze tests are written tests, which require the heritage speakers to be literate to some extent in their HL, alternative oral tasks, as the one we propose here, are necessary for children that are not formally instructed in their HL. In this sense, we propose that SRTs that include structures with different degrees of complexity, defined on the basis of previous studies, as described in this section, may complement the existing array of language assessment tools for EP, being adequate to assess monolingual and bilingual children’s development (Gaillard & Tremblay 2016).

3. The study

3.1. Research questions

The overall aim of this study is to assess the effectiveness of a SRT, developed for EP following the LITMUS procedures, in testing the language abilities of monolingual children and child heritage speakers of EP who grow up in a German-dominant environment. To achieve this overall aim, we ask the following five research questions:

Research Question 1 (RQ1): Do the children’s accuracy scores reflect the different levels of difficulty, defined on the basis of previous research on the acquisition of EP? That is, do children indeed score differently on the structures belonging to the three levels of difficulty, as shown for SRTs developed for other languages (Ibrahim & Hamann 2017, Komeili et al. 2020, Marinis & Armon-Lotem 2015, among others)?

Research Question 2 (RQ2): Do monolingual and bilingual children perform alike in the SRT or do child heritage speakers of EP perform lower than age-matched monolingual peers? If we observe differences, are these modulated by the different degrees of complexity of the tested structures? If so, which structures stand out as being particularly difficult for heritage bilingual children?

Research Question 3 (RQ3): Does the children's age at testing predict children's accuracy scores? If so, is the age effect visible in each complexity level and for both groups?

Research Question 4 (RQ4): Regarding the bilingual children, how do language experience variables, such as cumulative amount of HL exposure and richness of the HL input, influence performance of the child HSs in the task, when controlled for age and complexity level of the syntactic structures?

Research Question 5 (RQ5): Which scoring method, target structure (TS) or target structure & grammaticality (TS&Gra), is more fine-grained in capturing performance differences in the monolingual and bilingual groups?

3.2. Participants

Participants were 68 typically developing child speakers of EP between the ages of six and ten years. The children were divided into two groups: EP monolinguals and EP-German bilinguals. The monolingual group comprised 43 children (25 female) with a mean age of 8.05 years ($SD = 1.38$). All these children were born in Portugal and were being raised in a monolingual environment, as reported by their caregivers. As for the bilingual group, whose sociolinguistic data were collected by means of a detailed parental questionnaire (Correia & Flores 2021), it included 25 child heritage speakers of EP (12 female) with a mean age of 8.60 ($SD = 1.29$) who were living in Germany at the time of testing. Twenty children were born in Germany and the remaining five were born in Portugal, having immigrated to the host country before the age of 36 months. All the children had been exposed to EP from birth and half of them (13/25) had also been exposed to German since they were born. The remaining children started receiving systematic input in the majority language between the ages of 10 and 36 months. Furthermore, 21 out of the 25 children were being raised in families in which both parents spoke EP as a first language (L1); the remaining children belonged to households where just one of the parents had EP as an L1. With respect to the bilingual children's cumulative amount of HL exposure within the home⁴ (C.EXP.HL) and richness of the HL input⁵ (HL.RICHNESS), they ranged from 40.7 to 125 ($M = 88.2$; $SD = 22.5$) and from 0.05 to 0.63 ($M = 0.33$; $SD = 0.18$), respectively. Children's socioeconomic status (SES) was assessed based on the mothers' education level, which ranged from 1 (i.e., basic education/first cycle) to 5 (i.e., higher education) in both groups of speakers (monolinguals: $Mdn = 4$; $IQR = 1$; bilinguals: $Mdn = 4$, $IQR = 2$). No significant between-group differences were found regarding SES ($U = 483.5$, $p = .466$).

Participation in the study was voluntary and, prior to the study, a written parental informed consent was obtained for all participants. The study was approved by the Ethics Committee for Research in Social and Human Science of the University of Minho (CEICSH 004/2019 ADENDA1).

⁴In the parental questionnaire, the caregiver estimated, on a 5-point scale, the frequency of HL use from each household member to the child for several age periods of the child's life (i.e., from birth to current age). The scale was: *Never* (0%) = 0; *Seldom* (25%) = 0.25; *Sometimes* (50%) = 0.50; *Usually* (75%) = 0.75; *Always* (100%) = 1. For each age period, the quantity of language exposure within the home was calculated by totaling the responses and then dividing the resulting value by the number of household members, except for the child him-/herself and siblings under the age of two years. With these data, the quantity of HL input for each year of the child's life was determined and converted into months of exposure, which were then summed up to obtain the child's total amount of HL exposure in months over time within the household (for further details on calculation method, see Correia & Flores 2021).

⁵Richness of the HL input was calculated based on the child's language experience in four domains of his/her daily life, namely: (i) frequency of weekly contact with the HL by means of media-related activities (e.g., watching television, reading books, etc.); (ii) participation in leisure and sociocultural activities promoted by the Portuguese migrant community; (iii) HL use in the host country when among EP-speaking friends; and (iv) attendance of Portuguese classes. Points were given to each domain and then summed up and divided by the maximum possible score. The resulting value may range from 0 to 1; the closer the value is to 1, the greater the diversity and complexity of the HL input the child receives on a daily basis (for further details on calculation method, see Correia & Flores 2021).

3.3. Materials: the SRT-EP

The current SRT-EP was built on the first version of the task which was developed within the COST ACTION IS0804 and included 56 sentences targeting grammatical structures such as: simple sentences with periphrastic future and simple past, long passives, accusative clitics in enclisis and proclisis contexts, object *wh*-question with *who* and *which*, coordinate clauses, finite and non-finite complete clauses, and relative clauses (see Table A in the Online Supplementary Materials).

All these structures were maintained in the new version of the task, but some modifications were made in order to accommodate additional grammatical structures (i.e., SVO sentences with the subjunctive, short passives, and conditional clauses)⁶ and to solve some issues that were found to be problematic when the first version was piloted (e.g., the use of proper nouns in several stimuli). The main alterations involved: (i) reducing/expanding the number of sentences per target structure, so that each one would have the same number of stimuli (i.e., four sentences); (ii) substituting proper nouns for common nouns (e.g., *Maria* for *menina* 'girl'); (iii) avoiding contexts in which morphophonological properties would hinder the perception of the clitic. For the current version of the task, the stimuli were also pre-recorded by two adult monolingual speakers of EP (male and female) and incorporated in a PowerPoint presentation.

Moreover, to ensure that the lexical items included in the task would not be too difficult for the youngest children, we asked 11 primary school teachers to estimate on a 5-point scale, for every content word of the task ($n = 144$), the relative amount of 6-year-old typically developing children who knew its meaning. The scale was: *none* (0%) = 1, *a few* (+/- 25%) = 2, *half of them* (+/- 50%) = 3, *many* (+/- 75%) = 4, *all of them* (100%) = 5. The reasoning behind this methodological approach was that, due to their daily contact with children within the same age range of this study's participants, primary school teachers would provide more reliable information on the children's actual knowledge of the task's lexical items than some of the available EP lexical datasets such as ESCOLEX (Soares et al. 2014) or Portulex (Teixeira & Castro 2007). The latter databases are based on school textbooks, which may not cover the lexical knowledge of EP children who are starting to be literate. The outcome of the rating task revealed that the content words included in the SRT are suitable for children as young as six years old, since the medians of the ratings provided by primary school teachers ranged from 4 to 5 (see Table B in the Online Supplementary Materials for measures of central tendency and dispersion of each lexical item).

The resulting set of task stimuli comprised two practice items and 60 test items (four items per structure), with sentences varying from 10 to 18 syllables in length (see Table 1). The setting-up of the task was done in two steps. In the first step and in line with the literature on the syntactic development in EP (see Section 2.3), each target syntactic structure was assigned one of three increasing complexity levels, which ranged from C1 (less complex) to C3 (more complex).

In the second step and taking into account their assigned complexity level, the structures were grouped into three blocks and pseudo-randomized so that the children would not be consecutively presented with two items of the same target structure or the same lexical items. Care was also taken to present children with stimuli produced by a female and a male voice in an alternating way. Furthermore, bearing in mind the huge individual differences that typically characterize the language proficiency of bilingual speakers in their heritage language and in order to prevent feelings of frustration when performing the task, the first block included structures from level C1 and C2, and the second and third blocks included structures from level C2 and C3, as shown in Table 2.

3.4. Procedure and scoring

Participants were tested individually in a quiet room at their homes or, as was the case for some bilingual children, at the schools where they were attending Portuguese classes. The SRT was

⁶The initial version was piloted with pre-schoolers. Since the present version targeted older children, we believed it would be important to include a set of more demanding structures.

Table 1. Contents of the SRT-EP.

Structure type	Subtype	Example	Length (# syllables)
indicative	periphrastic future	Os meninos vão comer as maçãs no jardim. 'The boys are going to eat the apples in the garden.'	11-13
	past	A prima pintou o desenho com os lápis. 'The cousin colored the drawing with the pencils.'	12-13
complement clause	non-finite	A menina quer comer um gelado na praia. 'The girl wants to eat an ice cream on the beach.'	13-14
subjunctive	present	Talvez o pai esteja no trabalho amanhã. 'Maybe the father will be at work tomorrow.'	14
	past	Talvez a menina estivesse com medo do cão. 'Maybe the girl was afraid of the dog.'	15-16
coordinate clause	copulative	O homem tocou tambor e a mulher cantou na festa. 'The man played the drum and the woman sang at the party.'	16-17
	adversative	O pai falou alto, mas o bebé não acordou. 'The father spoke loudly, but the baby didn't wake up.'	14-15
passive	short	O rapaz foi assaltado na rua. 'The boy was robbed on the street.'	11-13
		O pirata foi avistado na ilha. 'The pirate was sighted on the island.'	
wh-question	who-object	Quem é que o polícia prendeu na estrada? 'Who did the policeman arrest on the street?'	13-14
complement clause	finite	O menino disse que brinca com carrinhos grandes. 'The boy said that he plays with big cars.'	14-15
relative clause	subject	Tu viste o rapaz que desenhou a mulher. 'You saw the boy that drew the woman.'	11-13
accusative clitic	enclisis	O dono da loja fechou-a depois do almoço. 'The owner of the store closed it after lunch.'	14-16
	proclisis	O gato do avô só o arranhou na perna. 'The grandfather's cat only scratched him on the leg.'	14-16
passive	long	O peixe foi comido pela baleia no mar. 'The fish was eaten by the whale in the sea.'	13-14
		O neto foi visto pelo avô no parque. 'The grandson was seen by the grandfather in the park.'	
relative clause	object	Tu viste a mulher que o rapaz desenhou. 'You saw the woman that the boy drew.'	11-13
wh-question	which-object	Que animal é que a cobra comeu ontem? 'Which animal did the snake eat yesterday?'	13-15
conditional clause	hypothetical or unreal	O menino apanharia a bola, se fosse mais alto. 'The boy would catch the ball if he were taller.'	14-18
		Se o pai fosse rico, compraria um carro. 'If the father were rich, he would buy a car.'	

Abbreviation: SRT-EP = Sentence Repetition Task-European Portuguese.

Table 2. Distribution of items by blocks in the SRT-EP.

Block 1: 20 sentences		Block 2: 24 sentences		Block 3: 16 sentences	
indicative	periphrastic future	accusative clitic	enclisis	relative clause	subject
subjunctive	past	passive	proclisis	wh-question	object
	present		short		which-object
coordinate clause	past	wh-question	long	conditional clause	hypothetical
complement clause	non-finite		who-object		
		complement clause	finite		

Abbreviation: Sentence Repetition Task-European Portuguese.

administered to each child by means of a PowerPoint presentation with embedded audio files. During the task administration, participants sat in front of a laptop computer and listened to the stimuli through headphones. They were instructed to listen to the sentences attentively and to repeat exactly what they heard. In order to familiarise children with the task, children were presented, prior to the experimental stimuli, with a practice session with two training sentences. In the actual experimental

task, children were allowed to listen to each stimulus only once unless some noise or interruption prevented them from hearing it. Children were orally praised from time to time irrespective of how successful their repetitions were. Self-corrections were allowed but only the final response was scored, whether it was correct or incorrect. Children's responses were audio recorded and later orthographically transcribed. The scoring was made offline based on the recorded audio files.⁷

The scoring procedure adopted for the present study included two measures: *target structure* (TS) and *target structure & grammaticality* (TS&Gra). The first measure is similar to the fifth scoring scheme (i.e., *sentence structure score*) presented by Marinis & Armon-Lotem (2015). According to this scoring method, a score of 1 is given if the child produces the structure that is targeted independently of whether (s)he makes changes in other parts of the sentence [see (10a)] with a number mismatch in the NP, which is not the targeted structure). A score of 0 is awarded if the child either makes an error in the targeted structure, e.g., omission of the accusative clitic in sentences targeting accusative clitics (see [10b]), or does not produce it and substitutes it with another structure⁸, e.g., using the indicative mood instead of the subjunctive one [see (10c)].

- (10) a. *A prima pintou o desenhos com o lápis.
 the cousin coloured the.M.SG drawings. M.PL with the pencils
 'The cousin coloured the drawings with the pencils.'
 target-like: A prima pintou o desenho com os lápis.
- b. *O gato do avô só [-] arranhou na perna.
 the cat of.the grandfather only [-] scratched on.the leg
 'The grandfather's cat only scratched [him] on the leg.'
 target-like: O gato do avô só o arranhou na perna.
- c. *Talvez a menina tem medo do cão.
 maybe the girl has fear of.the dog
 'Maybe the girl is afraid of the dog.'
 target-like: Talvez a menina estivesse com medo do cão.

As for the second measure, *target structure & grammaticality*, it is a combination of the previous one with the measure *grammaticality* proposed by Marinis & Armon-Lotem (2015). According to this

⁷The transcribed responses were scored by the first author and checked by the co-authors.

⁸Note that children were not penalized for replacing the conditional with the imperfect indicative in the conditional clauses, for example:

Se	a menina	arrumasse	o quarto,	ia	ao parque.
if	the girl	tidied	the room,	would.go	to.the park

'If the girl tidied the room, she would go to the park.'

target-like: Se a menina arrumasse o quarto, iria ao parque.

composite scoring scheme, a score of 1 is given only if the targeted structure is produced and the sentence has no grammatical errors despite the changes the child makes in it [see (11)]; otherwise, a score of 0 is awarded. Thus, with this scoring method, we aim not only at assessing whether the different syntactic structures are part of the children’s internal grammar, but also at evaluating children’s global grammatical/language proficiency.⁹

- (11) A menina sonhou que tem um chapéu azul.
the girl dreamt that has a hat blue
‘The girl dreamt that she has a blue hat.’
target-like: A menina sonhou que tinha um chapéu azul.

3.5. Statistical analyses

Statistical analyses were conducted in R (version 4.1.3, Core Team 2022), and data visualization was performed using the packages *ggplot2* (version 3.3.5; Wickham 2016) and *sjPlot* (version 2.8.10; Lüdtke 2021). The internal consistency of the SRT-EP was assessed with Cronbach’s alpha reliability coefficients, which were computed by means of the *alpha()* function of the *psych* package (version 2.2.5; Revelle 2022). To address the research questions, we fitted two sets of generalized linear mixed-effects models (GLMMs) to the data (see Table 3, for an overview) using the *lme4* package (version 1.1.29; Bates et al. 2015). Set A addressed the research questions focused on the two groups of EP-speaking children (RQ1–RQ3) whereas Set B addressed the research questions focused on the role that language experience variables play in the performance of the bilingual group (RQ4). RQ5 derives from the results of both sets of models.

Regarding Set A, it is composed of two subsets of analyses. Subset A1 comprised two models—for one of the models, the dependent variable was accuracy under the measure TS and, for the other one, it was accuracy under the measure TS&Gra. Predictors were *age* (in

Table 3. GLMMs in Sets A and B: overview of the dependent variables (DVs) and predictors.

Set	Subset	Model	DVs	Predictors
A	A1	1	TS	age (in months)
		2	TS&Gra	level of complexity (C1/C2/C3) group (monolingual/bilingual) two-way and three-way interactions
	A2	1	TS (C1)	age (in months)
		2	TS&Gra (C1)	group (monolingual/bilingual)
		3	TS (C2)	two-way interaction
		4	TS&Gra (C2)	
		5	TS (C3)	
		6	TS&Gra (C3)	
	B	1	TS	age (in months)
		2	TS&Gra	level of complexity (C1/C2/C3) cumulative amount of HL exposure richness of the HL input

Abbreviations: GLMMs = generalized linear mixed-effects models; TS = target structure; TS&Gra = target structure & grammaticality; HL = heritage language.

⁹In both scoring schemes, allowances were given for common phonological reductions (e.g., *tivesse* for *estivesse*), and expansions (e.g., *por a* for *pela*).

months), *level of complexity* (C1/C2/C3), *group* (monolingual/bilingual), and their interactions. Subset A2 included two models—one for *TS* and another one for *TS&Gra* - per participants' performance in each level of complexity. Predictors were *age* (in months) and *group* (monolingual/bilingual) as well as their interaction. With respect to Set B, it comprised two models. The dependent variables were *TS*, for one of the models, and *TS&Gra*, for the other one. Predictors were *cumulative amount of HL exposure within the home* (C.EXP.HL) and *richness of the HL input* (HL.RICHNESS) as well as *age* (in months) and *level of complexity* (C1/C2/C3). In each set of GLMMs, the dependent variables *TS* and *TS&Gra* were binomial categorical (two levels: 0 = inaccurate; 1 = accurate). Running a similar model for each measure enabled us to investigate whether differences in scoring method affect children's results in the SRT. Moreover, including the variables *age* and *level of complexity* in the models of Set B allowed us to assess the role of C.EXP.HL and HL.RICHNESS while taking into account the children's age and sentences' level of complexity. Before running the GLMMs, the continuous variables *age*, C.EXP.HL and HL.RICHNESS were scaled using the *scale()* function from the base package with *center* and *scale* set to *TRUE*. As for the categorical variables, they were coded as follows: the two-level factor *group* was coded using sum contrast coding (-.50/.50) and the three-level factor *level of complexity* was coded using simple contrast coding¹⁰ (UCLA Statistical Consulting 2011). We specified random intercepts for *participants* and *items* (i.e., *sentences*). Predicted probabilities derived from the models were obtained using the *ggpredict()* function from the *ggeffects* package (version 1.1.2; Lüdtke 2018).

4. Results

We will start by presenting the results of the reliability analyses aimed at assessing the internal consistency of the task. Then we will show the descriptive statistics regarding not only children's accuracy rates on the whole task, but also their accuracy rates broken down by level of complexity and grammatical structure. Finally, we will present the results that emerged from the two sets of GLMMs.

4.1. Internal consistency of the SRT-EP

In order to assess the internal consistency of the task, we conducted two sets of Cronbach's alpha reliability analyses on the items of the SRT: one set per measure (*TS* and *TS&Gra*). Within each set, Cronbach's alpha was calculated not only for the whole sample (i.e., monolingual and bilingual children together), but also for each group of speakers separately. As shown in Table 4, the internal consistency was excellent under the two scoring methods for the full sample and each subsample.

4.2. Performance in the SRT-EP

4.2.1. Descriptive statistics

Figure 1 depicts the descriptive data of each group's overall performance on the SRT by measure (i.e., *TS* and *TS&Gra*). A closer look at the boxplots shows that children's accuracy on both measures tended to be higher and less variable in the monolingual group than in the bilingual one. Furthermore, we observe lower accuracy rates in the measure *TS&Gra*, compared to the *TS* one.

Figure 2, in turn, illustrates each group's performance across the three levels of complexity, also by measure. A visual inspection of the data shows that accuracy rates tend to decrease as the level of

¹⁰When using simple contrast coding, the reference level is coded as -1/3 and the level that it is being compared to is coded as 2/3. For example, regarding the variable *level of complexity*, if one is comparing level C1 to level C2 and level C1 is the reference level, then the coding is 2/3 for level C2 and -1/3 for all other levels. Similar to dummy contrast coding, simple contrast coding compares each level to the reference level. However, in contrast to the former, the intercept of the latter represents the grand mean. Moreover, simple contrast coding has the advantage that "factors outside of interactions can be interpreted as main effects" (Tilmantine et al. 2021).

Table 4. Reliability statistics for both scoring methods.

	(n)	TS	TS&Gra
Monolinguals & Bilinguals	68	0.96	0.97
Monolinguals	43	0.90	0.91
EP-German bilinguals	25	0.97	0.96

Abbreviation: EP = European Portuguese; TS = target structure; TS & Gra = target structure & grammaticality.

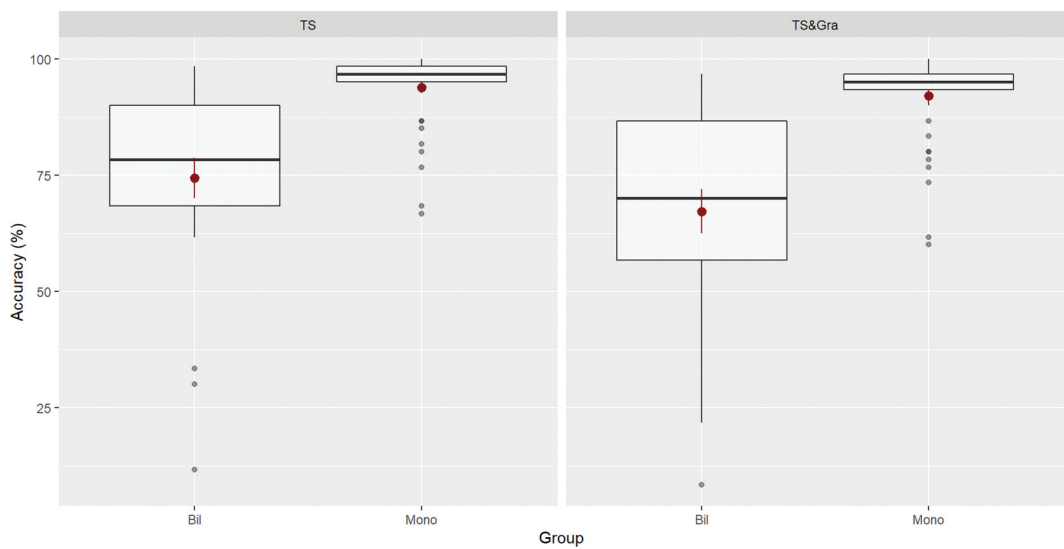


Figure 1. Overall performance (i.e., accuracy rate) on the SRT per scoring measure and group (the red dots and lines represent the mean and the standard error from the mean, respectively).

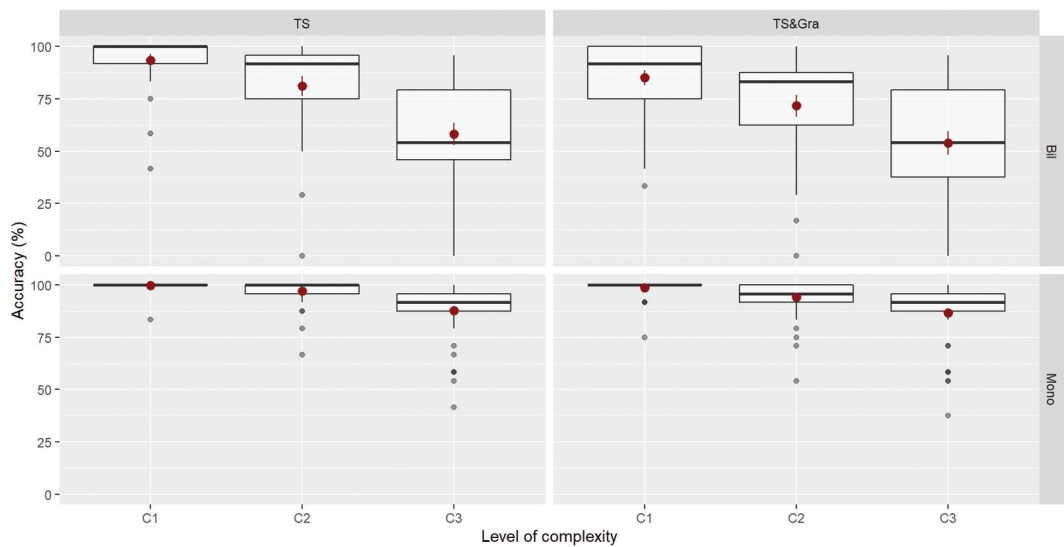


Figure 2. Performance (i.e., accuracy rate) on the SRT per level of complexity, scoring measure and group (the red dots and lines represent the mean and the standard error from the mean, respectively).

complexity of the structures included in the SRT increases and this occurs within each group and for the two scoring measures. Furthermore, the data display the same pattern observed for the children's overall performance on the SRT, that is, in each level of complexity, the monolingual group presents higher accuracy scores and less variability than the bilingual group.

The children's accuracy rates broken down by grammatical structure can be found in Table 5.

4.2.2. Generalized linear mixed-effects model analyses

4.2.2.1 Set A. Tables 6 and 7 report the outcome of the two generalized linear mixed-effects model analyses included in Subset A1, one for each scoring method. Regarding the measure *TS* (see Table 6), the overall findings of its model are as follows: (i) there is a significant effect of *level of complexity* when level C2 is being contrasted with level C3, but not when it is being compared to level C1 ($p = .059$), showing that the log-odds of providing an accurate answer decrease from level C2 to level C3, but they do not significantly increase from level C2 to level C1; (ii) there is a significant effect of *group*, with the positive estimates showing that belonging to the monolingual group increases the log-odds of providing an accurate answer; and (iii) there is also a significant effect of *age*, showing that the log-odds of providing an accurate answer increase with age. We found no significant two-way or three-way interactions.

Figure 3 depicts the predicted probabilities, based on the model, of a correct response under the measure *TS*. It shows that, in the bilingual group, a change in *age* from 1SD below the mean to 1SD above the mean leads to an increase in the probability of observing an accurate answer of: 2% (from 98% to 100%) with respect to the structures of level C1; of 9% (from 86% to 95%) regarding the structures of level C2; and of 28% (from 45% to 73%) in relation to structures of level C3. As for the monolingual group, moving from 1SD below the mean to 1SD above the mean in *age* does not lead to an increase/decrease in the probability of observing a correct response in the structures of level C1 since monolingual children tend to perform at ceiling in this level of the task; however, the probability of monolinguals producing an accurate answer increases by 1% (from 99% to 100%) among the items of level C2 and by 5% (from 94% to 99%) among the items of level C3.

Concerning the overall findings of the model focused on the measure *TS&Gra* (see Table 7), they are similar to the ones of the previous model, except for the fact that the effect of *level of complexity* is significant both when level C2 is being compared to level C3 as well as when it is being contrasted with level C1, revealing that the log-odds of producing an accurate response decrease from level C2 to level C3 but increase from level C2 to level C1.

Figure 4 illustrates the predicted probabilities of a correct response under the measure *TS&Gra*. It demonstrates that, in the bilingual group, a change in *age* from 1SD below the mean to 1SD above the mean leads to an increase in the probability of producing a correct response of: 2% (from 93% to 95%) regarding the structures of level C1; of 21% (from 68% to 89%) in relation to the structures of level C2; and of 25% (from 39% to 64%) with respect to the structures of level C3. Regarding the monolingual group, going from 1SD below the mean to 1SD above the mean in *age* increases the probability of observing an accurate answer by 1% (from 99% to 100%) among the items of level C1, by 3% (from 97% to 100%) among the items of level C2, and by 7% (from 92% to 99%) among the items of level C3.

As a final step in the analyses of Subset A1, we calculated the Coefficient of Discrimination of each model, using the *r2_tjur()* function of the *performance* package (version 0.9.0; Lüdtke et al. 2021). The outcome revealed that the fixed effects account for 53% of the variation in the dependent variable *TS* and for 50% of the variation in the dependent variable *TS&Gra*.

With respect to the analyses performed in Subset A2, aimed at assessing the effects of *group* and *age* in each complexity level, Tables 8 and 9 report the outcome of the GLMMs that were run for each measure per level of complexity. Regarding the measure *TS* (see Table 8), the overall findings of its models are as follows: (i) for level C1, the predictors *age*, *group*, and the two-way interaction *age*group* are non-significant, meaning that, among the items of level C1, the log-odds of observing a correct response do not change as a function of the children's age (for both monolinguals and bilinguals) nor as a function of the group they belong to; (ii) for levels C2 and C3, there is a significant effect of *age* and

Table 5. Performance (i.e., accuracy rate) on the SRT structures per scoring measure and group.

CLevel	Structure	Measure	Group	N	Mean	SE	SD	Min.	Max.
C1	CompNonFinite	TS	Bil	25	88.0	5.025	25.12	0	100
			Mono	43	100.0	0.000	0.00	100	100
		TS&Gra	Bil	25	81.0	5.260	26.30	0	100
	PeriFuture	TS	Mono	43	99.4	0.581	3.81	75	100
			Bil	25	98.0	1.384	6.92	75	100
		TS&Gra	Bil	25	89.0	3.253	16.27	50	100
	Preterite	TS	Mono	43	98.8	0.812	5.33	75	100
			Bil	25	94.0	2.986	14.93	50	100
		TS&Gra	Bil	25	85.0	4.082	20.41	25	100
		TS	Mono	43	97.7	1.825	11.97	25	100
			Bil	25	88.0	5.228	26.14	0	100
		TS&Gra	Bil	25	79.0	5.530	27.65	0	100
C2	CompFinite	TS	Mono	43	97.1	1.490	9.77	50	100
			Bil	25	79.0	5.530	27.65	0	100
		TS&Gra	Bil	25	79.0	5.530	27.65	0	100
	Coordinate	TS	Mono	43	95.3	1.907	12.51	50	100
			Bil	25	94.0	4.153	20.77	0	100
		TS&Gra	Bil	25	85.0	4.330	21.65	0	100
	PassiveShort	TS	Mono	43	92.4	2.569	16.85	25	100
			Bil	25	90.0	5.204	26.02	0	100
		TS&Gra	Bil	25	78.0	6.178	30.89	0	100
	RelativeSub	TS	Mono	43	98.3	0.983	6.44	75	100
			Bil	25	88.0	5.795	28.98	0	100
		TS&Gra	Bil	25	83.0	6.076	30.38	0	100
	Subjunctive	TS	Mono	43	98.3	1.288	8.44	50	100
			Bil	25	57.0	7.292	36.46	0	100
		TS&Gra	Bil	25	45.0	6.922	34.61	0	100
	WhoObject	TS	Mono	43	89.5	3.139	20.58	0	100
			Bil	25	70.0	7.773	38.86	0	100
		TS&Gra	Bil	25	60.0	8.898	44.49	0	100
		TS	Mono	43	90.7	2.634	17.27	25	100
			Bil	25	60.0	7.071	35.36	0	100
		TS&Gra	Bil	25	59.0	7.053	35.27	0	100
C3	Conditional	TS	Mono	43	94.2	2.862	18.77	0	100
			Bil	25	59.0	7.053	35.27	0	100
		TS&Gra	Bil	25	59.0	7.053	35.27	0	100
	Enclisis	TS	Mono	43	93.6	2.893	18.97	0	100
			Bil	25	37.0	6.144	30.72	0	100
		TS&Gra	Bil	25	32.0	5.867	29.33	0	75
	PassiveLong	TS	Mono	43	63.4	4.583	30.05	0	100
			Bil	25	68.0	7.433	37.17	0	100
		TS&Gra	Bil	25	63.0	7.789	38.94	0	100
	Proclisis	TS	Mono	43	97.1	1.237	8.11	75	100
			Bil	25	63.0	7.789	38.94	0	100
		TS&Gra	Bil	25	63.0	7.789	38.94	0	100
	RelativeObj	TS	Mono	43	97.1	1.237	8.11	75	100
			Bil	25	29.0	6.874	34.37	0	100
		TS&Gra	Bil	25	29.0	6.874	34.37	0	100
	WhichObject	TS	Mono	43	75.0	5.128	33.63	0	100
			Bil	25	29.0	6.874	34.37	0	100
		TS&Gra	Bil	25	29.0	6.874	34.37	0	100
		TS	Mono	43	75.0	5.128	33.63	0	100
			Bil	25	89.0	5.788	28.94	0	100
		TS&Gra	Bil	25	86.0	6.137	30.69	0	100
		TS	Mono	43	99.4	0.581	3.81	75	100
			Bil	25	86.0	6.137	30.69	0	100
		TS&Gra	Bil	25	86.0	6.137	30.69	0	100
		TS	Mono	43	99.4	0.581	3.81	75	100
			Bil	25	66.0	7.053	35.27	0	100
		TS&Gra	Bil	25	54.0	7.455	37.28	0	100
		TS	Mono	43	95.3	1.716	11.25	50	100
			Bil	25	54.0	7.455	37.28	0	100
		TS&Gra	Bil	25	54.0	7.455	37.28	0	100
		TS	Mono	43	90.7	2.499	16.39	50	100
			Bil	25	54.0	7.455	37.28	0	100
		TS&Gra	Bil	25	54.0	7.455	37.28	0	100

Note: SRT = Sentence Repetition Task; CLevel = level of complexity; SE = standard error; SD = standard deviation; Min. = minimum; Max. = maximum; TS = target structure; TS&Gra = target structure & grammaticality.

Table 6. Summary of the GLMM for the measure TS (Subset A1).^a

Predictors	TS				
	Estimate	SE	CI	z	p
(Intercept)	4.32	0.47	3.40 – 5.24	9.21	<0.001
Age in months	1.00	0.34	0.33 – 1.67	2.91	0.004
Complexity (C1 vs. C2)	2.21	1.17	-0.08 – 4.50	1.89	0.059
Complexity (C3 vs. C2)	-2.35	0.54	-3.40 – -1.29	-4.35	<0.001
Group (Mono vs. Bil)	3.71	0.81	2.13 – 5.29	4.60	<0.001
Age in months * Complexity (C1 vs. C2)	-0.05	0.78	-1.57 – 1.47	-0.06	0.951
Age in months * Complexity (C3 vs. C2)	-0.40	0.24	-0.88 – 0.07	-1.68	0.093
Age in months * Group (Mono vs. Bil)	0.67	0.68	-0.67 – 2.01	0.98	0.327
Complexity (C1 vs. C2) * Group (Mono vs. Bil)	0.08	2.02	-3.87 – 4.04	0.04	0.967
Complexity (C3 vs. C2) * Group (Mono vs. Bil)	-0.66	0.58	-1.80 – 0.48	-1.13	0.257
Age in months * Complexity (C1 vs. C2) * Group (Mono vs. Bil)	-0.39	1.55	-3.43 – 2.65	-0.25	0.802
Age in months * Complexity (C3 vs. C2) * Group (Mono vs. Bil)	-0.82	0.48	-1.76 – 0.12	-1.70	0.088

Note: ^aThe fixed effects, their estimates, standard errors (SE), confidence intervals (CI), statistic (z), and p-values are given. Formula in R for the model TS: TS ~ Age in months * Complexity * Group + (1|Participant) + (1|Sentence). The reference level is *bilingual* for the variable *group* and level C2 for the variable *level of complexity*.

Abbreviation: GLMMs = generalized linear mixed-effects models; TS = target structure.

Table 7. Summary of the GLMMs for the measure TS&Gra (Subset A1).^a

Predictors	TS&Gra				
	Estimate	SE	CI	z	p
(Intercept)	3.08	0.31	2.48 – 3.68	10.00	<0.001
Age in months	0.72	0.22	0.29 – 1.15	3.28	0.001
Complexity (C1 vs. C2)	1.59	0.64	0.33 – 2.85	2.48	0.013
Complexity (C3 vs. C2)	-1.29	0.45	-2.17 – -0.41	-2.87	0.004
Group (Mono vs. Bil)	3.29	0.46	2.39 – 4.18	7.17	<0.001
Age in months * Complexity (C1 vs. C2)	-0.22	0.30	-0.82 – 0.37	-0.73	0.464
Age in months * Complexity (C3 vs. C2)	-0.14	0.14	-0.42 – 0.13	-1.01	0.311
Age in months * Group (Mono vs. Bil)	0.52	0.44	-0.34 – 1.38	1.18	0.240
Complexity (C1 vs. C2) * Group (Mono vs. Bil)	0.34	0.73	-1.09 – 1.78	0.47	0.638
Complexity (C3 vs. C2) * Group (Mono vs. Bil)	0.10	0.31	-0.50 – 0.71	0.34	0.734
Age in months * Complexity (C1 vs. C2) * Group (Mono vs. Bil)	0.43	0.61	-0.76 – 1.62	0.72	0.475
Age in months * Complexity (C3 vs. C2) * Group (Mono vs. Bil)	-0.02	0.28	-0.57 – 0.52	-0.08	0.934

Note: ^aThe fixed effects, their estimates, standard errors (SE), confidence intervals (CI), statistic (z), and p-values are given. Formula in R for the model TS&Gra: TS_Gra ~ Age in months * Complexity * Group + (1|Participant) + (1|Sentence). The reference level is *bilingual* for the variable *group* and level C2 for the variable *level of complexity*. Abbreviations: GLMMs = generalized linear mixed-effects models; TS&Gra = target structure & grammaticality.

of *group*, with the positive estimates showing that, among the items of both level C2 and level C3, children perform better with increasing age and belonging to the monolingual group increases the log-odds of giving an accurate response. Furthermore, the lack of interaction between *age* and *group* indicates that the effect of age holds for both groups.

As for the measure *TS&Gra* (see Table 9), the findings of its models are similar to the ones found for the measure *TS*, except for the fact that there is a main effect of *group* among the items of level C1, meaning that the log-odds of providing a correct answer increase if children belong to the monolingual group.

The final step in this set of analyses also included the calculation of the Coefficient of Discrimination of each model, which revealed that, for the dependent variable *TS*, the fixed effects account for: 48% of the variation with the items of level C1, 46% of the variation with the items of level C2, and 54% of the variation with the items of level C3. As for the dependent variable *TS&Gra*, they account for: 33%, 44%, and 55% of the variation with the items of level C1, C2, and C3, respectively.

4.2.2.2 Set B. Table 10 and Table 11, in turn, report the outcome of the two generalized linear mixed-effects model analyses included in Set B, which includes only the bilingual group.



Figure 3. Measure TS: predicted probabilities of an accurate answer plotted by age, group and level of complexity.

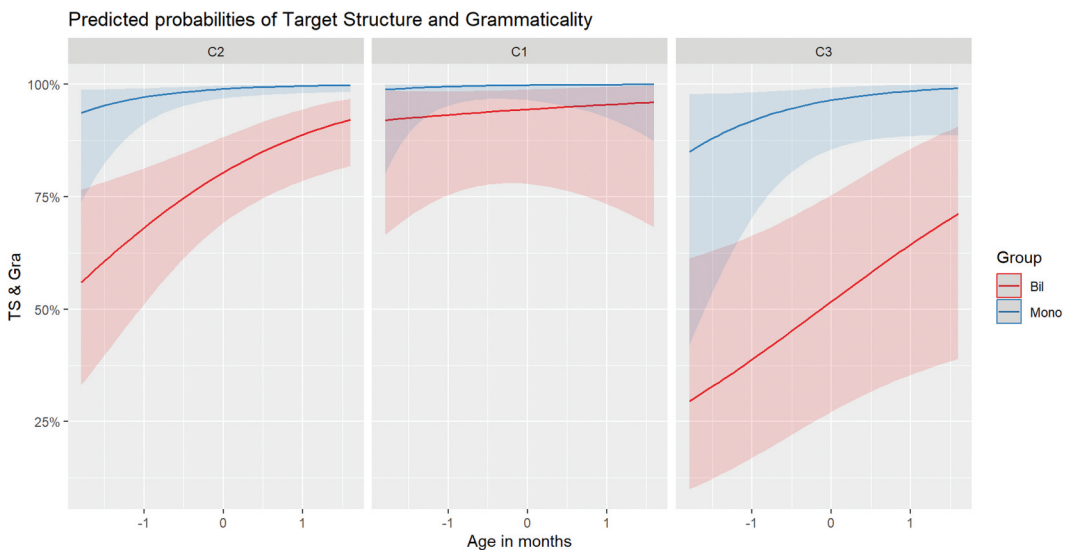


Figure 4. Measure TS&Gra: predicted probabilities of an accurate answer plotted by age, group and level of complexity.

Regarding the measure *TS* (see Table 10), the overall findings of its model are as follows: (i) there is a significant effect of *age and level of complexity* in the bilingual group even when language experience variables are taken into account, meaning that the log-odds of producing a correct answer not only increase with age, but they also decrease as the level of difficulty of the targeted structures increases (i.e., they increase from level C2 to level C1 and decrease from level C2 to level C3); and (ii) there is a significant effect of *the richness of the HL input*, with the positive estimate showing that being exposed to more diverse HL input increases the log-odds of giving an accurate response, even when age and level of complexity of the stimuli are controlled for. No significant effect of cumulative amount of HL exposure was found.

Table 8. Summary of the GLMMs for the measure TS (Subset A2).^a

Predictors (C1)	TS				
	Estimate	SE	CI	z	p
(Intercept)	10.03	3.06	4.04 – 16.03	3.28	0.001
Age in months	1.32	1.79	-2.19 – 4.83	0.74	0.460
Group (Mono vs. Bil)	4.63	4.61	-4.40 – 13.66	1.00	0.315
Age in months * Group (Mono vs. Bil)	0.76	3.57	-6.25 – 7.76	0.21	0.833
Predictors (C2)					
(Intercept)	4.50	0.52	3.49 – 5.52	8.68	<0.001
Age in months	1.37	0.37	0.64 – 2.09	3.69	<0.001
Group (Mono vs. Bil)	4.37	0.82	2.76 – 5.98	5.32	<0.001
Age in months * Group (Mono vs. Bil)	1.36	0.73	-0.07 – 2.79	1.86	0.062
Predictors (C3)					
(Intercept)	2.11	0.46	1.21 – 3.01	4.58	<0.001
Age in months	0.74	0.25	0.26 – 1.22	3.01	0.003
Group (Mono vs. Bil)	3.37	0.51	2.38 – 4.37	6.67	<0.001
Age in months * Group (Mono vs. Bil)	0.23	0.49	-0.73 – 1.18	0.46	0.644

Note: ^aThe fixed effects, their estimates, standard errors (SE), confidence intervals (CI), statistic (z), and p-values are given. Formula in R for the models: TS ~ Age in months * Group + (1|Participant) + (1|Sentence). The reference level is *bilingual* for the variable *group*. Abbreviations: GLMMs = generalized linear mixed-effects models; TS = target structure.

Table 9. Summary of the GLMMs for the measure TS&Gra (Subset A2).^a

Predictors (C1)	TS&Gra				
	Estimate	SE	CI	z	p
(Intercept)	4.59	0.66	3.30 – 5.88	6.97	<0.001
Age in months	0.79	0.43	-0.06 – 1.63	1.83	0.068
Group (Mono vs. Bil)	4.02	1.00	2.07 – 5.98	4.03	<0.001
Age in months * Group (Mono vs. Bil)	1.11	0.86	-0.58 – 2.80	1.29	0.198
Predictors (C2)					
(Intercept)	3.02	0.38	2.28 – 3.77	7.97	<0.001
Age in months	0.91	0.24	0.44 – 1.38	3.77	<0.001
Group (Mono vs. Bil)	3.37	0.51	2.37 – 4.36	6.62	<0.001
Age in months * Group (Mono vs. Bil)	0.61	0.48	-0.33 – 1.54	1.28	0.202
Predictors (C3)					
(Intercept)	1.83	0.45	0.95 – 2.70	4.10	<0.001
Age in months	0.71	0.24	0.24 – 1.18	2.94	0.003
Group (Mono vs. Bil)	3.45	0.49	2.48 – 4.41	6.99	<0.001
Age in months * Group (Mono vs. Bil)	0.31	0.48	-0.63 – 1.26	0.65	0.514

Note: ^aThe fixed effects, their estimates, standard errors (SE), confidence intervals (CI), statistic (z), and p-values are given. Formula in R for the models: TS&Gra ~ Age in months * Group + (1|Participant) + (1|Sentence). The reference level is *bilingual* for the variable *group*. Abbreviations: GLMMs = generalized linear mixed-effects models; TS&Gra = target structure & grammaticality.

Figure 5 displays the predicted probabilities, based on the model, of obtaining a correct response under the measure TS. Focusing our analysis on the language experience predictor that reached significance (i.e., *richness of the HL input*) and taking as reference mean-aged children, one can see that a change in HL.RICHNESS from 1SD below the mean to 1SD above the mean leads to an increase in the probability of observing an accurate answer of: 2% (from 98% to 100%) in relation to the structures of level C1; of 14% (from 84% to 98%) with respect to the items of level C2; and of 47% (from 37% to 84%) concerning the items of level C3.

As far as the overall findings of the model focused on the measure TS&Gra (see Table 11) are concerned, they are in line with the ones of the previous model, that is, there is a significant effect of *age*, *level of complexity*, and *richness of the HL input*. This means that the log-odds of observing a correct response increase with growing age and with greater exposure to more diverse and complex HL input, on the one hand, and decrease as the complexity level of the stimuli increases, on the other hand. The model did not find a significant effect of *cumulative amount of HL exposure*.

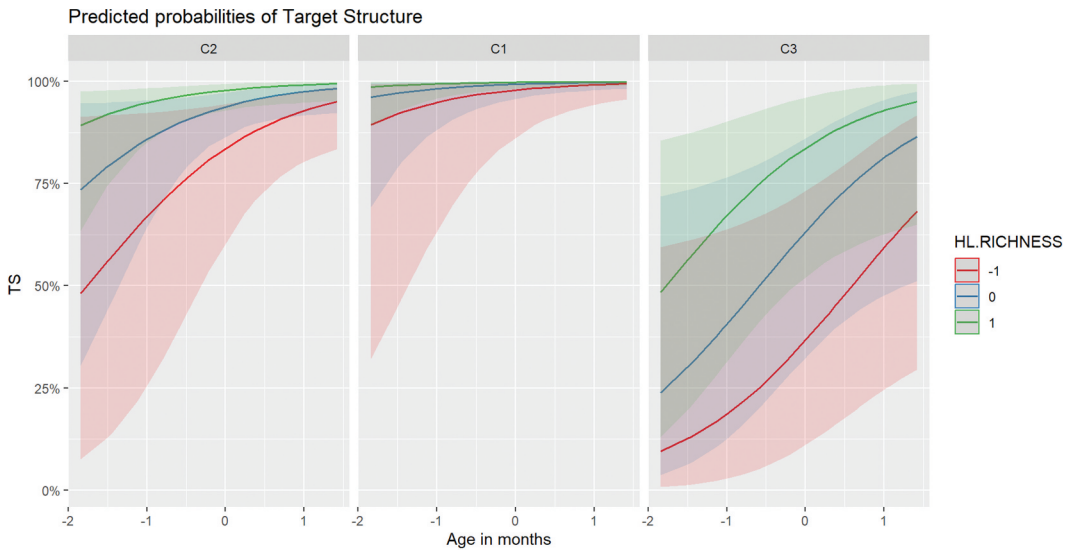


Figure 5. Measure TS (Set B): predicted probabilities of an accurate answer plotted by age, richness of the HL input (HL.RICHNESS) and level of complexity.

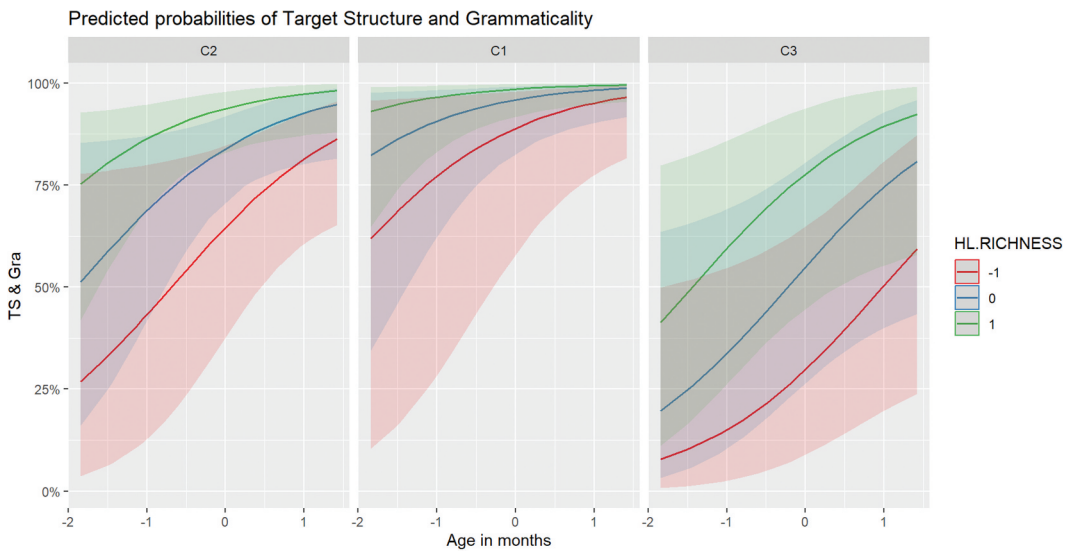


Figure 6. Measure TS&Gra (Set B): predicted probabilities of an accurate answer plotted by age, richness of the HL input (HL.RICHNESS) and level of complexity.

Figure 6 illustrates the predicted probabilities, derived from the model, of an accurate answer under the measure *TS&Gra*. Focusing again on the predictor *richness of the HL input* and also taking as reference the values for mean-aged children, one can see that moving 1SD below the mean to 1SD above the mean in *HL.RICHNESS* leads to an increase in the probability of observing a correct response: of 10% (from 89% to 99%) among the items of level C1; of 29% (from 65% to 94%) among the items of level C2; and of 48% (from 30% to 78%) among the items of level C3.

Finally, the results of the Coefficient of Discrimination of each model revealed that the fixed effects account for 55% of the variation in the dependent variable *TS* and for 51% of the variation in the dependent variable *TS&Gra*.

Table 10. Summary of the GLMMs for the measure TS (Set B).^a

Predictors	TS				
	Estimate	SE	CI	z	p
(Intercept)	2.73	0.43	1.89 – 3.58	6.36	<0.001
Age in months	0.93	0.46	0.02 – 1.83	2.00	0.045
Complexity (C1 vs. C2)	2.21	0.70	0.82 – 3.59	3.13	0.002
Complexity (C3 vs. C2)	-2.18	0.52	-3.20 – -1.16	-4.20	<0.001
C.EXP.HL	0.27	0.45	-0.62 – 1.15	0.58	0.559
HL.RICHNESS	1.09	0.45	0.22 – 1.97	2.44	0.015

Note. ^aThe fixed effects, their estimates, standard errors (SE), confidence intervals (CI), statistic (z), and p-values are given. Formula in R for the model TS: TS ~ Age in months + Complexity + C.EXP.HL + HL.RICHNESS + (1|Participant) + (1|Sentence). The reference level is level C2 for the variable *level of complexity*. Abbreviations: GLMMs = generalized linear mixed-effects models; TS = target structure.

Table 11. Summary of the GLMM for the measure TS&Gra (Set B).^a

Predictors	TS&Gra				
	Estimate	SE	CI	z	p
(Intercept)	1.67	0.39	0.91 – 2.43	4.32	<0.001
Age in months	0.87	0.43	0.04 – 1.71	2.04	0.041
Complexity (C1 vs. C2)	1.49	0.63	0.26 – 2.72	2.37	0.018
Complexity (C3 vs. C2)	-1.46	0.50	-2.43 – -0.49	-2.95	0.003
C EXP HL	0.17	0.42	-0.65 – 0.99	0.40	0.686
HL RICHNESS	1.06	0.42	0.24 – 1.87	2.55	0.011

Note: ^aThe fixed effects, their estimates, standard errors (SE), confidence intervals (CI), statistic (z), and p-values are given. Formula in R for the model TS: TS&Gra ~ Age in months + Complexity + C.EXP.HL + HL.RICHNESS + (1|Participant) + (1|Sentence). The reference level is level C2 for the variable *level of complexity*. Abbreviations: GLMMs = generalized linear mixed-effects models; TS&Gra = target structure & grammaticality.

5. Discussion

In the present study, we tested EP monolingual and EP-German bilingual children's language proficiency with a SRT developed for EP on the basis of the LITMUS-SRT version. The results give us interesting insights into the use of this instrument to capture the particularities of the children's language proficiency as well as into the way individual differences in the performance of the child HSs are modulated by language experience variables. We discuss the results following the research questions listed in Section 3.1.

The first question (RQ1) asked whether, in general, EP-speaking children score differently on the structures belonging to the different levels of complexity, as shown extensively for other languages (Ibrahim & Hamann 2017, Komeili et al. 2020, Marinis & Armon-Lotem 2015, among others). The results indicate that this is clearly the case (see Figure 2). Both the monolingual and the bilingual children perform almost at ceiling with respect to structures classified as belonging to the lowest level of complexity (C1), in particular if the scoring method *TS* is applied. The scores decrease in level C2, particularly in the bilingual group, and are even lower in the highest level of complexity (C3). The GLMMs of Subset A1 identify, for both scoring methods, the level of complexity as significant predictor, with the probability of scoring lower decreasing significantly from level C2 to level C3. The difference between C1 and C2 is only captured by the more rigorous *TS&Gra* scoring method. This indicates that the SRT-EP indeed captures the particularities of EP children's language development, being sensitive to different degrees of complexity of the linguistic properties.

Furthermore, the SRT-EP captures not only the differences reflected by the different degrees of complexity, but also the inter-group dissimilarities. Answering the second research question (RQ2), we can, in fact, conclude that the child heritage speakers of EP perform lower than their age-matched monolingual peers, mainly when they are assessed by means of the measure *TS&Gra*. The GLMMs in

Subset A1 identify *group* as significant predictor of children's accuracy scores, with a higher probability of monolingual children giving an accurate answer compared to bilingual children. The between-group differences are further supported by the GLMMs in Subset A2, which found a group effect within each complexity level under the measure TS&Gra and within levels C2 and C3 under the measure TS. These outcomes are, thus, in line with those of many other studies that applied SRTs to test bilingual children's language proficiency and that have found that, in the HL, bilinguals were outperformed by their monolingual peers (e.g., Haman et al. 2017, Meir et al. 2016, and others).

A detailed look at the mean accuracy rates of the various structures corroborates the results obtained in the analyses of Subset A2 and gives us a more thorough understanding not only of the between-group differences, but also of the structures that stand out as being particularly challenging for the heritage bilingual children. As shown in Table 5, in the monolingual group, the accuracy rates for non-finite complement clauses (CompNonFinite), simple clauses with periphrastic future (PeriFuture), and with Simple Past tense (Preterite) lie between 98.8% and 100% for the TS scoring method and between 97.7% and 99.4% if the TS&Gra method is applied. This indicates that these structures have been fully acquired by most monolinguals. With respect to their bilingual counterparts, despite the fact that these speakers present slightly lower accuracy rates and more variation in their scores (i.e., TS: 88-98%; TS&Gra: 81-89%), one may also conclude that these structures have also been acquired by the majority of the bilingual children, providing evidence that EP-speaking children acquire these grammatical structures in pre-school age, regardless of growing up in a monolingual or in a bilingual context. The picture changes somehow when we look at the structures belonging to the level C2. While the mean scores in the monolingual group are still very high for all structures and for both scoring methods (above 89.5%, even though lower than at C1), the bilingual children reveal difficulties with respect to various structures belonging to level C2. Particularly challenging for the child heritage speakers of EP are the sentences with the subjunctive and 'who-object' wh-questions. A look at the C3 structures indicates that these are even more demanding for the bilingual children, with monolinguals also performing lower in this level of complexity. Sentences with accusative clitics, both in enclisis and in proclisis position, are the most challenging for the bilingual children, displaying the lowest accuracy rates (29% and 37%). This coincides with the structures displaying the lowest accuracy rates in the monolingual group, whose accuracy rates range from 63.4% to 75%. Conditional clauses, which-object questions and long passives, in contrast, displayed results above 90% in the monolingual group, but proved to be difficult structures for the bilinguals, whose highest accuracy rate in these structures is 68%. Against our expectation, object relatives were not problematic for either group, with bilinguals ranging from 86% to 89%, and monolinguals being almost at ceiling (99.4%). There are several factors that may explain why object wh-questions were more difficult than object relatives for the heritage bilinguals: object wh-questions in Portuguese require that the C position is filled with the grammaticalized "é que" expression. This is quite language-specific. Relative clauses, on the contrary, do not have this language-specific feature. A qualitative analysis of children's non-target productions shows that, although some children make errors that correspond to theta-role reversal and are, thus, attributable to problems with intervention, most errors are related to difficulties in the production of the "é que" expression that fills the C position. Therefore, we believe that the difference between relatives and wh-questions in children's productions may be due to late mastery of this specific grammatical feature.

In sum, the analysis of the various accuracy scores per structure confirms findings from previous studies which showed that heritage bilinguals demonstrate protracted development mainly in late acquired (and more complex) structures (see Tsimpli 2014). For child heritage speakers of EP, these structures encompass clauses with subjunctive mood (as already shown by Flores et al. 2017, 2019 for mood selection in complement clauses), and with clitics in various positions, as shown in several studies with bilinguals (Costa et al. 2016, Flores & Barbosa 2014; Flores et al. 2016; Nardelli & Lobo 2018; Rinke & Flores 2014; Tomaz et al. 2019, 2020). Additionally, our results show that some structures that are already acquired at this age for most monolingual children can still be problematic for bilinguals: these include conditional clauses, long passives and which-object questions.

Our third research question (RQ3) asked whether the children's current age predicts their accuracy scores, thus showing language development with increasing age, even in primary school-aged children. Similar to previous findings from studies on bilingual and monolingual children, which were conducted with either SRTs (e.g., Komeili et al. 2020, Paradis et al. 2021, Taha et al. 2021, Theodorou et al. 2017) or other grammatical assessment tasks (e.g., Cadório et al. 2014, Flores et al. 2017, Torregrossa et al. 2023), the results from the current study reveal that the investigated monolingual and bilingual EP-speaking children perform better with growing age, that is, as shown in the GLMMs of Subset A1, the log-odds of giving an accurate answer increase as children grow older (but see, e.g., Fleckstein et al. 2018 for different outcomes). Moreover, the GLMMs of Subset A2, focused on the children's accuracy by complexity level, together with the predicted probabilities derived from the GLMMs of Subset A1 (plotted by age, group and level of complexity) indicate, on the one hand, that the effect of age becomes more visible as the level of difficulty increases, and, on the other hand, that it is larger for the bilingual children (see Figure 3 and Figure 4). In a nutshell, these results not only suggest a similar qualitative trend in grammatical development for both monolingual and bilingual children as a function of age, but they also support the assumption that a heritage language develops through the whole childhood, as long as heritage children maintain regular contact with their heritage language—even in cases of gradual input reduction due to schooling in the majority language (Rodina et al. 2023). Corroborating the results from previous studies (Flores & Barbosa 2014; Flores et al. 2017, 2020; Torregrossa et al. 2023), our results on EP-speaking heritage bilinguals do not support the idea of stagnation or language attrition in heritage language development in childhood.

With respect to our fourth research question (RQ4), it investigated the role that language experience variables, such as *cumulative amount of HL exposure* and *richness of the HL input*, play in bilingual children's performance in the task. The analyses performed in Set B reveal both expected and unexpected results. On the one hand, they show that, for both measures (*TS* and *TS&Gra*), the greater the diversity and complexity of the HL input the bilingual children are exposed to through certain activities and interactions, the more accurate are their responses in the SRT. This outcome is in line with previous research that shows a positive relationship between richness of the HL environment and HSS' performance not only in experimental tasks aimed at assessing HL morphosyntactic (e.g., Ibrahim et al. 2020) and narrative abilities (e.g., Jia & Paradis 2015), but also in tasks aimed at evaluating HL vocabulary knowledge (e.g., Pham & Tipton 2018). On the other hand, our results further demonstrate that *cumulative amount of HL exposure*, a language experience variable that has been found to affect HSS' performance in tasks evaluating either HL morphosyntactic (e.g., Mitrofanova et al. 2018, Rodina & Westergaard 2017) or lexical knowledge (e.g., Haman et al. 2017, Tao et al. 2021), does not emerge as a significant predictor of the bilingual children's accuracy once variables such as *age*, *complexity level*, and *HL richness* are taken into account. Taken together, the outcomes of the GLMMs focused on the potential predictors of individual differences in the bilingual group indicate that contact with diverse sources of HL input, that is, media, HL-speaking friends, sociocultural activities in the immigrant community, and HL formal instruction, is a better predictor of the syntactic abilities of child HSS of EP than cumulative amount of HL exposure within the home. This could be explained by the fact that HL input at home is usually restricted to informal spoken registers and to a small number of interlocutors—sometimes, to just one of the parents. As stated by Polinsky (2015:10), the HL input HSS receive at home “is not representative of the speech of the entire native-speaking population, nor does it cover all the possible contexts in which a language can be used.” Moreover, depending on the background of the migrant families, the verbal productions of their members may exhibit traces of language attrition and/or other effects of prolonged language contact with the societal language (Pascual y Cabo & Rothman 2012). Bearing this in mind, it may come to no surprise that HSS who are exposed to more diverse, complex and formal HL input in their daily lives perform better in the SRT than their counterparts whose language exposure comes from fewer sources of HL input.

Finally, we asked which scoring method, *TS* or *TS&Gra*, is more fine-grained in capturing performance differences in the monolingual and bilingual groups. Results show that children have lower accuracy scores

if we apply the *TS&Gra* scoring method, which seems to indicate that this scoring method is more conservative than the *TS* one. This is particularly relevant in the bilingual group, where the *TS&Gra* scoring is more fine-grained at identifying children with lower proficiency levels, who produce ungrammatical sentences. We would, thus, recommend using this scoring method in future studies.

To close this discussion, the results of the Cronbach's alpha reliability analyses on the items of the SRT reveal high internal consistency of the task (see Table 4) both for the whole sample (i.e., monolingual and bilingual children together), as well as for each group separately. This confirms that the SRT presented in this study is internally consistent and suitable to test EP-speaking children's language proficiency, at least in the age span between 6 and 10 years. In the bilingual group, it allows us to capture developmental differences between the bilingual children regarding their HL, caused by age and by diverse input experiences. In the monolingual group, even though the accuracy rates are high and many children perform at ceiling with respect to less complex structures, it still captures performance differences caused by age, thus showing that this task is useful to evaluate children's knowledge of more complex structures. Testing of younger monolingual children will presumably reveal lower accuracy rates.

6. Conclusion

Overall, our results show that the EP version of the SRT may be used as an assessment tool for both monolingual and bilingual children's language abilities and is able to discriminate between different levels of proficiency according to the complexity level of the structures. In accordance with previous studies, our research was also able to identify areas that are more vulnerable in bilingual development compared to the monolingual one and areas that stabilize early. In addition, the three complexity levels, defined for this task on the basis of previous findings from studies on monolingual and bilingual acquisition of EP, are proven to accurately capture the particularities of the linguistic system of EP.

Acknowledgments

The authors are grateful to the children and parents who participated in this study. We would also like to thank Camões, Institute of Cooperation and Language I.P. and all the individuals who helped us recruiting participants both in Portugal and in Germany, as well as Jacopo Torregrossa for helpful advice during data analysis. We extend special thanks to João Costa and Fernanda Pratas, who, in collaboration with Maria Lobo, developed the first version of this task within the COST ACTION IS0804. Finally, we would also like to thank not only Theo Marinis and all members of the aforementioned project for their ideas, support, and fruitful discussions, but also the participants of the Workshop LITMUS Sentence Repetition Task, held at University of Konstanz in 2019, for their valuable comments on the current study.

Disclosure Statement

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This study was partially supported by a PhD Studentship (SFRH/BD/138397/2018, for Liliana Correia) from the Portuguese Foundation for Science and Technology (FCT) through national funds (Portuguese Republic|Ministry for Science, Technology and Higher Education) and the European Social Fund (Human Capital Operational Program – Portugal 2020; Regional Operational Program – Norte 2020). Another part of this research (Maria Lobo) was supported by the Portuguese Foundation for Science and Technology, I.P. (FCT) as part of the project UIDB/LIN/03213/2020; 10.54499/UIDB/03213/2020 and UIDP/LIN/03213/2020; 10.54499/UIDP/03213/2020 – Linguistics Research Centre of NOVA University Lisbon (CLUNL). Cristina Flores' contribution was funded by FCT – Fundação para a Ciência e Tecnologia through the project <https://doi.org/10.54499/UIDB/00305/2020>.

ORCID

Liliana Correia  <http://orcid.org/0000-0003-1045-6119>

Maria Lobo  <http://orcid.org/0000-0002-3290-3381>

Cristina Flores  <http://orcid.org/0000-0001-5629-9556>

Data Availability Statement

Due to the nature of this research, participants of this study did not agree for their data to be shared publicly.

References

- Adani, Flavia, Heather K. J. van der Lely, Matteo Forgiarini & Maria Teresa Guasti. 2010. Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua* 120(9). 2148–2166. doi:10.1016/j.lingua.2010.03.018.
- Agostinho, Celina. 2014. *The acquisition of control in European Portuguese complement clauses*. Lisbon: University of Lisbon dissertation.
- Agostinho, Celina. 2020. *The acquisition of the verbal passive in European Portuguese*. Barcelona: Universitat Autònoma de Barcelona dissertation.
- Aguiar, Joana. 2017. *Mecanismos de conexão frásica: a importância das variáveis sociais* [Phrasal connection mechanisms: the importance of social variables]. Unpublished Doctoral Dissertation. Universidade do Minho, Portugal.
- Almeida, Laetitia de, Sandrine Ferré, Eléonore Morin, Philippe Prévost, Christophe dos Santos, Laurie Tuller, Racha Zebib & Marie-Anne Barthez. 2017. Identification of bilingual children with specific language impairment in France. *Linguistic Approaches to Bilingualism* 7(3–4). 331–358. doi:10.1075/lab.15019.alm.
- Andreou, Maria, Jacopo Torregrossa & Christiane Bongartz. 2021. Sentence Repetition Task as a measure of language dominance. In Danielle Dionne and Lee-Ann Vidal Covas (eds.), *Proceedings of the 45th annual Boston University Conference on Language Development*, 14–25. Somerville, MA: Cascadia Press.
- Armon-Lotem, Sharon, Jan de Jong & Natalia Meir, eds. 2015. *Assessing Multilingual Children*. Disentangling Bilingualism from Language Impairment. Bristol: Multilingual Matters. doi:10.21832/9781783093137.
- Baião, Vera & Maria Lobo. 2014. Aquisição de interrogativas preposicionadas no português europeu [Acquisition of prepositional interrogatives in European Portuguese]. In António Moreno, Maria de Fátima Silva, Isabel Falé, Isabel Pereira & João Veloso (eds.), *Textos Seleccionados XXIX Encontro Nacional da APL*, 57–70. Coimbra: APL.
- Barbosa, Pilar, Cristina Flores & Cátia Pereira. 2018. On subject realization in infinitival complements of causative and perceptual verbs in European Portuguese. Evidence from monolingual and bilingual speakers. In Pedro Guijarro-Fuentes and Alejandro Cuza (eds.), *Language acquisition and contact in the Iberian Peninsula*, 125–158. Berlin, Boston: Mouton de Gruyter. doi:10.1515/9781501509988-006.
- Batalha, Joana, Maria Lobo, Antónia Estrela & Bruna Bragança. 2022. Desenvolvimento sintático em produções escritas de crianças de 1.º ciclo [Syntactic development in written productions of primary school children]. *Revista da Associação Portuguesa de Linguística* 9.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. doi:10.18637/jss.v067.i01.
- Bentea, Anamaria. 2017. Comprehension of wh-questions in child Romanian: A case about case and lexical restriction. In Jiyoung Choi, Hamida Demirdache, Oana Lungu & Laurence Voeltzel (eds.), *Language Acquisition and Development: Proceedings of GALA 2015*, 1–19. Newcastle upon Tyne: CSP.
- Bentea, Anamaria, Stephanie Durrleman & Luigi Rizzi. 2016. Refining intervention: The acquisition of featural relations in object A-bar dependencies. *Lingua* 169. 21–41. doi:10.1016/j.lingua.2015.10.001.
- Bever, Thomas G. 1970. The cognitive basis for linguistic structures. In John R. Hayes (ed.), *Cognition and the development of language*, 279–362. New York: John Wiley and Sons, Inc.
- Blume, Maria & Barbara C. Lust. 2017. *Research methods in language acquisition: Principles, procedures and practices*. Berlin: Walter de Gruyter.
- Borer, Hagit & Kenneth Wexler. 1987. The maturation of syntax. In Thomas Roeper and Edwin Williams (eds.), *Parameter setting*, 123–172. Dordrecht: Springer. doi:10.1007/978-94-009-3727-7_6.
- Cadório, Inês, Maria Lousada, Melina Aparici & Andreia Hall. 2014. Assessment of morphosyntactic development in European Portuguese-speaking children. *Folia phoniatrica et logopaedica* 66(6). 251–257. doi:10.1159/000368333.
- Cerejeira, Joana. 2009. *Aquisição de interrogativas de sujeito e de objecto em português europeu* [Acquisition of subject and object interrogatives in European Portuguese]. Lisbon: NOVA University of Lisbon dissertation.
- Chiat, Shula, Sharon Armon-Lotem, Theodoros Marinis, Kamila Polišínská, Penny Roy & Belinda Seeff-Gabriel. 2013. The potential of sentence imitation tasks for assessment of language abilities in sequential bilingual children. In

- Virginia C. Mueller Gathercole (ed.), *Bilinguals and assessment: State of the art guide to issues and solutions from around the world*, 56–89. Bristol: Multilingual Matters.
- Corrêa, Letícia M., Marina R. A. Augusto & João C. de Lima-Júnior. 2017. Passivas [Passives]. In Maria João Freitas and Ana Lúcia Santos (eds.), *Aquisição de língua materna e não materna: Questões gerais e dados do português*, 201–224. Berlin: Language Science Press.
- Correia, Liliana, & Cristina Flores. 2021. Questionário sociolinguístico parental para famílias emigrantes bilingues (QuesFEB): uma ferramenta de recolha de dados sociolinguísticos de crianças falantes de herança [Sociolinguistic parental questionnaire for bilingual migrant families (QuesFEB): a tool for collecting sociolinguistic data from heritage-speaking children]. *Linguística. Revista de Estudos Linguísticos da Universidade do Porto* 16. 75–102. <https://ojs.letras.up.pt/index.php/EL/article/view/11021/10077>
- Costa, Ana Luísa. 2010. *Estruturas contrastivas: desenvolvimento do conhecimento explícito e da competência de escrita* [Contrastive structures: the development of explicit knowledge and writing skills]. Lisbon: University of Lisbon dissertation.
- Costa, Ana, Nélia Alexandre, Ana Lúcia Santos & Nuno Soares. 2008. Efeitos de modelização no input: o caso da aquisição de conectores [Input modeling effects: the case of the acquisition of connectors]. In Sónia Frota and Ana Lúcia Santos (eds.), *Textos Seleccionados do XXIII ENAPL*, 131–142. Lisbon: Colibri.
- Costa, João, Alexandra Fiéis & Maria Lobo. 2015. Input variability and late acquisition: Clitic misplacement in European Portuguese. *Lingua* 161. 10–26. doi:10.1016/j.lingua.2014.05.009.
- Costa, João & Maria Lobo. 2007. Clitic omission, null objects or both in the acquisition of European Portuguese? In Sergio Baauw, Frank Drijkoningen & Manuela Pinto (eds.), *Romance Languages and Linguistic Theory 2005. Selected Proceedings of Going Romance*, 59–71. Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/cilt.291.06cos.
- Costa, João, Maria Lobo & Fernanda Pratas. 2016. Clitic production by Portuguese and Capeverdean children: omission in bilingualism. *Probus* 28(2). 271–291. doi: 10.1515/probus-2014-0015.
- Costa, João, Maria Lobo & Carolina Silva. 2011. Subject-object asymmetries in the acquisition of Portuguese relative clauses: Adults vs. children. *Lingua* 121(6). 1083–1100. doi:10.1016/j.lingua.2011.02.001.
- De Houwer, Annick. 2009. *An introduction to bilingual development*. Bristol: Multilingual Matters. doi:10.21832/9781847691705.
- Delage, Hélène & Laurie Tuller. 2010. Evolution of syntactic complexity and avoidance strategies in children and adolescents with mild-to-moderate hearing loss. In João Costa, Ana Castro, Maria Lobo & Fernanda Pratas (eds.), *Language Acquisition and Development: Proceedings of GALA 2009*, 107–120. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.
- Durrleman, Stephanie & Anamaria Bentea. 2021. Locality in the acquisition of object A'-dependencies: Insights from French. *Glossa: A Journal of General Linguistics* 6(1). 106. doi:10.16995/glossa.5876.
- Ebert, Kerry Danahy. 2014. Role of auditory non-verbal working memory in sentence repetition for bilingual children with primary language impairment. *International Journal of Language & Communication Disorders* 49. 631–36.
- Erlam, Rosemary. 2006. Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics* 27(3). 464–491. doi:10.1093/applin/aml001.
- Estrela, Antónia. 2013. *A aquisição da estrutura passiva em Português Europeu* [The acquisition of the passive in European Portuguese]. Lisbon: NOVA University Lisbon dissertation.
- Fleckstein, Alice, Philippe Prevost, Laurice Tuller, Eva Sizaret & Rasha Zebib. 2018. How to identify SLI in bilingual children: A study on sentence repetition in French. *Language Acquisition* 25(1). 85–101. doi:10.1080/10489223.2016.1192635.
- Flores, Cristina & Pilar Barbosa. 2014. When reduced input leads to delayed acquisition: A study on the acquisition of clitic placement by Portuguese heritage speakers. *International Journal of Bilingualism* 18(3). 304–325. doi:10.1177/1367006912448124.
- Flores, Cristina, Pilar Barbosa & Manuela Casa Nova. 2016. A closer look at cross-linguistic influence in the acquisition of Portuguese as a heritage language. In Sambor Grucza, Magdalena Olpińska-Szkielko & Piotr Romanowski (eds.), *Bilingual landscape of the contemporary world*, 75–94. Frankfurt/Main, London, New York: Peter Lang Verlag.
- Flores, Cristina, Ana Lúcia Santos, Letícia Almeida, Alice Jesus & Rui Marques. 2019. Portuguese as a heritage language in contact with German and French: A comparative study on the acquisition of verbal mood. In Ingo Feldhausen, Martin Elsig, Imme Kuchenbrandt & Mareike Neuhaus (eds.), *Romance languages and linguistic theory 15. Selected papers from 'Going Romance' 30, Frankfurt*, 36–52. Amsterdam: John Benjamins. doi:10.1075/rllt.15.02flo.
- Flores, Cristina, Ana Lúcia Santos, Alice Jesus & Rui Marques. 2017. Age and input effects in the acquisition of mood in Heritage Portuguese. *Journal of Child Language* 44(4). 795–828. doi:10.1017/S0305000916000222.
- Flores, Cristina, Esther Rinke & Aldona Sopata. 2020. Acquiring the distribution of null and overt direct objects in European Portuguese. *Journal of Portuguese Linguistics* 19(1). 5. doi:10.5334/jpl.239.
- Fox, Danny & Yosef Grodzinsky. 1998. Children's passive: A view from the by-phrase. *Linguistic Inquiry* 29(2). 311–332.
- Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativised relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1). 67–88. doi:10.1016/j.lingua.2008.09.002.
- Friedmann, Naama & Rama Novogrodsky. 2011. Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of SLI. *Lingua* 121(3). 367–382. doi: 10.1016/j.lingua.2010.10.004.

- Gaillard, Stéphanie & Annie Tremblay. 2016. Linguistic proficiency assessment in second language acquisition research: The Elicited Imitation Task. *Language Learning* 66(2). 419–447. doi:10.1111/lang.12157.
- Gonçalves, Fernanda, Paula Guerreiro & Maria João Freitas. 2011. *O conhecimento da língua: percursos de desenvolvimento* [Language knowledge: development paths]. Lisbon: DGIDC/Ministério da Educação.
- Gordon, Peter C., Randall Hendrick & Marcus Johnson. 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 27, 1411–1423.
- Gordon, Peter C., Randall Hendrick & Marcus Johnson. 2004. Effects of noun phrase type on sentence complexity. *Journal of Memory and Language* 51 (1), 97–114.
- Gordon, Peter C. & Matthew W. Lowder. 2012. Complex sentence processing: A review of theoretical perspectives on the comprehension of relative clauses. *Language and Linguistics Compass* 6/7. 403–415.
- Grillo, Nino. 2009. Generalized Minimality: Feature impoverishment and comprehension deficits in agrammatism. *Lingua* 119(10). 1426–1443. doi:10.1016/j.lingua.2008.04.003.
- Hamann, Cornelia, Solveig Chilla, Lina Abed Ibrahim & István Fekete. 2020. Language assessment tools for Arabic-speaking heritage and refugee children in Germany. *Applied Psycholinguistics* 41(6). 1375–1414. doi:10.1017/S0142176420000399.
- Haman, Ewa, Zofia Wodniecka, Marta Marecka, Jakub Szewczyk, Marta Bialecka-Pikul, Agnieszka Otwinowska, Karolina Mieszkowska, Magdalena Łuniewska, Joanna Kołak, Aneta Mięksiz, Agnieszka Kacprzak, Natalia Banasik & Małgorzata Forys-Nogala. 2017. How does L1 and L2 exposure impact L1 performance in bilingual children? Evidence from Polish-English migrants to the United Kingdom. *Frontiers in Psychology* 8. 1444. doi:10.3389/fpsyg.2017.01444.
- Hirsch, Christopher & Ken Wexler. 2006. By the way, children don't know by. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development*, 249–261. Somerville, MA: Cascadia Press.
- Hoff, Erika, Cynthia Core, Silvia Place, Rosario Rumiche, Melissa Señor & Marisol Parra. 2012. Dual language exposure and early bilingual development. *Journal of Child Language* 39. 1–27. doi:10.1017/S0305000910000759
- Hoff, Erika, Rosario Rumiche, Andrea Burridge, Krystal M. Ribot & Stephanie N. Welsh. 2014. Expressive vocabulary development in children from bilingual and monolingual homes: A longitudinal study from two to four years. *Early Childhood Research Quarterly* 29(4). 433–444. doi:10.1016/j.ecresq.2014.04.012.
- Ibrahim, Lina Abed & Cornelia Hamann. 2017. Bilingual Arabic-German and Turkish-German children with and without specific language impairment: Comparing performance in sentence and Nonword Repetition Tasks. In Maria LaMendola and Jennifer Scott (eds.), *Proceedings of the 41st Annual Boston University Conference on Language Development*, 1–17. Somerville, MA: Cascadia Press.
- Ibrahim, Abed Lina, Cornelia Hamann & István Fekete. 2020. Language assessment of bilingual Arabic-German heritage and refugee children: Comparing performance on LITMUS Repetition Tasks. *Proceedings of BUCLD 44*. 1–17. Somerville, MA: Cascadia Press.
- Jesus, Alice, Rui Marques & Ana Lúcia Santos. 2019. Semantic features in the acquisition of mood in European Portuguese. *Language Acquisition* 26(3). 302–338. doi:10.1080/10489223.2019.1570203.
- Jia, Ruiting & Johanne Paradis. 2015. The use of referring expressions in narratives by Mandarin heritage language children and the role of language environment factors in predicting individual differences. *Bilingualism* 18. 737–735
- Komeili, Mariam, Theodoros Marinis, Parvaneh Tavakoli & Yalda Kazemi. 2020. Sentence repetition in Farsi-English bilingual children. *Journal of the European Second Language Association* 4(1). 1–12. doi:10.22599/jesla.55.
- Lewis, Richard L., Shravan Vasishth & Julie A. Van Dyke. 2006. Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences* 10(10), 447–454.
- Lima Júnior, João Claudio, Marina Augusto & Letícia Sicuro Corrêa. 2018. Discrimination of passive predicates by Brazilian Portuguese-speaking children. In Anna Gavarró (ed.), *On the acquisition of the syntax of romance*, 31–56. Amsterdam: John Benjamins.
- Lira, R., L. M. Sicuro Corrêa & Maria Lobo. in press. Animacy and schooling in the production of null objects and accusative clitics in the acquisition of two varieties of Portuguese. In Pilar Barbosa, Cristina Flores, A. C. Silva & E.-M. Roessler (eds.), *Null objects from a crosslinguistic and developmental perspective*. Series Language Faculty and Beyond. John Benjamins.
- Lobo, Maria & Carla Soares-Jesel. 2017. Interrogativas, relativas e clivadas. In Maria João Freitas and Ana Lúcia Santos (eds.), *A aquisição de língua materna e não materna. Questões gerais e dados do Português* [The acquisition of mother tongue and non-mother tongue. General questions and data for Portuguese], 225–248. Berlin: Language Science Press.
- Lüdecke, Daniel. 2018. ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software* 3(26).772. doi:10.21105/joss.00772.
- Lüdecke, Daniel. 2021. sjPlot: Data visualization for statistics in social science. R package version 2.8.10. <https://strengjacker.github.io/sjPlot/>
- Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, Philip Waggoner & Dominique Makowski. 2021. Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software* 6(60). 3139. doi: 10.21105/joss.03139.

- Lust, Barbara, Suzanne Flynn & Claire Foley. 1996. What children know about what they say: elicited imitation as a research method for assessing children's syntax. In Dana McDaniel, Cecile McKee & Helen Smith Cairns (eds.), *Methods for Assessing Children's Syntax*, 55–76. Cambridge, MA/London: The MIT Press.
- Maratsos, Michael, Dana E.C. Fox, Judith A. Becker & Mary Anne Chalkley. 1985. Semantic restrictions on children's passives. *Cognition* 19(2). 167–191. doi:[10.1016/0010-0277\(85\)90017-4](https://doi.org/10.1016/0010-0277(85)90017-4).
- Marinis, Theodoros & Sharon Armon-Lotem. 2015. Sentence repetition. In Sharon Armon-Lotem, Jan de Jong & Natalia Meir (eds.), *Assessing multilingual children. disentangling bilingualism from language impairment*, 95–122. Bristol: Multilingual Matters.
- Martins, Ana Maria. 2013. Posição dos pronomes pessoais clíticos. In Eduardo P. Raposo, Maria Fernanda Bacelar do Nascimento, Maria Antónia Mota, Luísa Segura & Amália Mendes (eds.), *Gramática do Português*, 2231–2302. Lisboa: Fundação Calouste Gulbenkian.
- Meir, Natalia, Joel Walters & Sharon Armon-Lotem. 2016. Disentangling SLI and bilingualism using sentence repetition tasks: The impact of L1 and L2 properties. *International Journal of Bilingualism* 20(4). 421–452. doi:[10.1177/1367006915609240](https://doi.org/10.1177/1367006915609240).
- Mendes, Ana, Elisabete Afonso, Marisa Lousada & Fátima Andrade. 2014. *Teste de Linguagem ALPE* [ALPE Language Test]. Aveiro: Edubox.
- Mitrofanova, Natalia, Yulia Rodina, Olga Urek & Marit Westergaard. 2018. Bilinguals' sensitivity to grammatical gender cues in Russian: The role of cumulative input, proficiency, and dominance. *Frontiers in Psychology* 9. 1–20. doi:[10.3389/fpsyg.2018.01894](https://doi.org/10.3389/fpsyg.2018.01894).
- Montrul, Silvina. 2016. *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
- Nardelli, Marina & Maria Lobo. 2018. Clitic omission in bilingual Portuguese-Spanish acquisition. In Anna Gavarró (ed.), *On the acquisition of the syntax of romance*, 119–145. Amsterdam: John Benjamins. doi: [10.1075/lald.62.06nar](https://doi.org/10.1075/lald.62.06nar).
- Novogrodsky, Rama & Naama Friedmann. 2006. The production of relative clauses in SLI: A window to the nature of the impairment. *Advances in Speech-Language Pathology* 8(4). 364–375. doi:[10.1080/14417040600919496](https://doi.org/10.1080/14417040600919496).
- Okura, Eve & Deryle Lonsdale. 2012. Working memory's meager involvement in sentence repetition tests. In *Proceedings of the Annual Meeting of the Cognitive Science Society* 34. 2131–2137.
- Paradis, Johanne. 2023. Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language*. 1–25. doi:[10.1017/S0305000922000708](https://doi.org/10.1017/S0305000922000708).
- Paradis, Johanne, Adriana Soto-Corominas, Evangelia Daskalaki, Xi Chen & Alexandra Gottardo. 2021. Morphosyntactic development in first generation Arabic–English children: The effect of cognitive, age, and input factors over time and across languages. *Languages* 6. 51. doi:[10.3390/languages6010051](https://doi.org/10.3390/languages6010051).
- Pascual y Cabo, Diego & Jason Rothman. 2012. The (Il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics* 33(4). 450–455. <https://doi.org/10.1093/applin/ams037>
- Peristeri, Eleni, Maria Andreou & Ianthi M. Tsimpli. 2017. Syntactic and story structure complexity in the narratives of high- and low-language ability children with autism spectrum disorder. *Frontiers in Psychology* 8. 2027. doi:[10.3389/fpsyg.2017.02027](https://doi.org/10.3389/fpsyg.2017.02027).
- Pham, Giang & Timothy Tipton. 2018. Internal and external factors that support children's minority first language and English. *Language, Speech, and Hearing Services in Schools* 49. 595–606.
- Polinsky, Maria. 2015. Heritage languages and their speakers: State of the field, challenges, perspectives for future work, and methodologies. *Zeitschrift für Fremdsprachwissenschaft* 26. 7–27.
- Prentza, Alexandra, Dionysios Tafiadis, Vasiliki Chondrogianni & Ianthi-Maria Tsimpli. 2022. Validation of a Greek Sentence Repetition Task with typically developing monolingual and bilingual children. *Journal of Psycholinguistic Research* 51(2). 373–395. doi:[10.1007/s10936-022-09853-z](https://doi.org/10.1007/s10936-022-09853-z).
- R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Raposo, Eduardo. 1986. On the null object construction in European Portuguese. In Osvaldo Jaeggli & Carmen Silva-Corvalán (eds.), *Studies in Romance Linguistics*, 373–390. Dordrecht: Foris.
- Revelle, William. 2022. *psych: Procedures for personality and psychological research*. Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych>
- Riches, Nick. 2012. Sentence repetition in children with specific language impairment: an investigation of underlying mechanisms. *International Journal of Language & Communication Disorders* 47. 499–510. <https://doi.org/10.1111/j.1460-6984.2012.00158.x>
- Rinke, Esther & Cristina Flores. 2014. Morphosyntactic knowledge of clitics by Portuguese heritage bilinguals. *Bilingualism: Language and Cognition* 17(4). 681–699. doi:[10.1017/S136672891300076X](https://doi.org/10.1017/S136672891300076X).
- Rodina, Yulia, Alexandra Bogoyavlenskaya, Natalia Mitrofanova & Marit Westergaard. 2023. Russian heritage language development in narrative contexts: Evidence from pre- and primary-school children in Norway, Germany, and the UK. *Frontiers in Psychology* 14. 1101995. doi: [10.3389/fpsyg.2023.1101995](https://doi.org/10.3389/fpsyg.2023.1101995).
- Rodina, Yulia, Tanja Kupisch, Natalia Meir, Natalia Mitrofanova, Olga Urek & Marit Westergaard. 2020. Internal and external factors in heritage language acquisition: Evidence from heritage Russian in Israel, Germany, Norway, Latvia and the United Kingdom. *Frontiers in Education* 5. 1–17. doi: [10.3389/feduc.2020.00020](https://doi.org/10.3389/feduc.2020.00020).

- Rodina, Yulia & Marit Westergaard. 2017. Grammatical gender in bilingual Norwegian–Russian acquisition: The role of input and transparency. *Bilingualism: Language and Cognition* 20(1). 197–214. doi:10.1017/S1366728915000668.
- Santos, Ana Lúcia. 2017. Alguns aspetos da aquisição de orações subordinadas completivas [Some aspects of the acquisition of complement clauses]. In Maria João Freitas and Ana Lúcia Santos (eds.), *Aquisição de língua materna e não materna: Questões gerais e dados do português*, 249–273. Berlin: Language Science Press.
- Santos, Ana Lúcia, Anabela Gonçalves & Nina Hyams. 2016. Aspects of the acquisition of object control and ECM-Type verbs in European Portuguese. *Language Acquisition* 23(3). 199–233. doi:10.1080/10489223.2015.1067320.
- Scheidnes, Maureen, Laurice Tuller & Hélène Delage. 2009. L1 transfer versus computational complexity in adult L2 French: evidence from a comparison with deaf L1 French Learners. In Jean Crawford, Koichi Otaki & Masahiko Takahashi (eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)*, 241–252. Somerville, MA: Cascadia Proceedings Project.
- Schönström, Krister & Peter C. Hauser. 2022. The sentence repetition task as a measure of sign language proficiency. *Applied Psycholinguistics* 43(1). 157–175. doi:10.1017/S0142716421000436.
- Sim-Sim, Inês. 2006. *Avaliação da linguagem oral: Um contributo para o conhecimento do desenvolvimento linguístico das crianças portuguesas* [Oral language assessment: A contribution to the knowledge of the linguistic development of Portuguese children]. Lisbon: Fundação Calouste Gulbenkian.
- Snyder, William & Nina Hyams. 2015. Minimality effects in children’s passives. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond*, 343–368. Amsterdam: John Benjamins.
- Soares, Ana Paula, José Carlos Medeiros, Alberto Simões, João Machado, Ana Costa, Álvaro Iriarte, José João de Almeida, Ana P. Pinheiro & Montserrat Comesaña. 2014. ESCOLEX: A grade-level lexical database from European Portuguese elementary to middle school textbooks. *Behavior Research Methods* 46(1). 240–253. doi:10.3758/s13428-013-0350-1.
- Soares, Carla. 2006. *La syntaxe de la périphérie gauche en portugais européen et son acquisition* [The syntax of the left periphery in European Portuguese and its acquisition]. Paris: Université Paris 8 dissertation.
- Spada, Nina, Julie Li-Ju Shiu & Yasuyo Tomita. 2015. Validating an Elicited Imitation Task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning* 65(3). 723–751. doi:10.1111/lang.12129.
- Taha, Juhayna, Vesna Stojanovic & Emma Pagnamenta. 2021. Sentence repetition as a clinical marker of developmental language disorder: Evidence from Arabic. *Journal of Speech, Language, and Hearing Research* 64(12). 4876–4899. doi:10.1044/2021_JSLHR-21-00244.
- Tao, Lily, Qing Cai, Tamar Gollan. 2021. Effects of cumulative language exposure on heritage and majority language skills. Spanish and Mandarin heritage speakers in the USA. *Linguistic Approaches to Bilingualism* 11(2). 168–191. <https://doi.org/10.1075/lab.18044.tao>
- Teixeira, C. & São Luís Castro. 2007. *Portulex: Base de dados lexical de manuais do 1º ao 4º anos de escolaridade* [Lexical database of textbooks for the 1st to 4th years of schooling]. Laboratório de Fala da Faculdade de Psicologia da Universidade do Porto.
- Theodorou, Elena, Maria Kambanaros & Kleanthes K. Grohmann. 2017. Sentence repetition as a tool for screening morphosyntactic abilities of bilingual children with SLI. *Frontiers in Psychology* 8. 2104. doi:10.3389/fpsyg.2017.02104
- Tilmatine, Mesian, Ferdy Hubers & Florian Hintz. 2021. Exploring Individual Differences in Recognizing Idiomatic Expressions in Context. *Journal of Cognition* 4(1). 37. doi:10.5334/joc.183.
- Tomaz, Margarida, Maria Lobo, Ana Madeira, Carla Soares-Jesel & Stéphanie Vaz. 2019. Omissão e colocação de clíticos por crianças bilingues Português-Francês [Effects of language dominance and timing of formal language exposure on the production of clitic pronouns by Portuguese/French bilingual children]. *Revista da Associação Portuguesa de Linguística* 5. 385–412. doi: 10.26334/2183-9077/rapln5ano2019a25.
- Tomaz, Margarida, Maria Lobo, Ana Madeira, Carla Soares-Jesel & Stéphanie Vaz. 2020. Efeitos de dominância linguística e de tempo de exposição formal à língua na produção de pronomes clíticos por crianças bilingues português/francês. *Revista da Associação Portuguesa de Linguística* 7. 350–367. doi:10.26334/2183-9077/rapln7ano2020a21.
- Torregrossa, Jacopo, Cristina Flores & Esther Rinke. 2023. What modulates the acquisition of difficult structures in a heritage language? A study on Portuguese in contact with French, German and Italian. *Bilingualism: Language and Cognition* 26(1). 179–192. doi:10.1017/S1366728922000438.
- Tsimpli, Ianthi Maria. 2014. Early, late or very late? Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism* 4(3). 283–313. doi:10.1075/lab.4.3.01tsi.
- UCLA Statistical Consulting Group. 2011. *R library: Contrast coding systems for categorical variables*. Los Angeles, CA: UCLA Statistical Consulting Group. <https://stats.idre.ucla.edu/r/library/r-library-contrast-coding-systems-for-categorical-variables> (August 2022).
- Unsworth, Sharon. 2013. Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition* 16(1). 86–110. doi:10.1017/S1366728912000284.
- Unsworth, Sharon. 2015. Amount of exposure as a proxy for dominance in bilingual language acquisition. In Carmen Silva-Corvalan & Jeanine Treffers-Daller (eds.), *Language dominance in bilinguals: Issues of measurement and operationalization*, 156–173. Cambridge, UK: Cambridge University Press.
- Varlokosta, Spyridoula, Adriana Belletti, João Costa, Naama Friedmann, Anna Gavarró, Kleanthes K. Grohmann, Maria Teresa Guasti, et al. 2016. A cross-linguistic study of the acquisition of clitic and pronoun production. *Language Acquisition* 23(1). 1–26. doi:10.1080/10489223.2015.1028628.

- Vasconcelos, Manuela. 1993. Relative clauses acquisition and experimental research: A study with Portuguese children. In Isabel Hub Faria and Maria João Freitas (eds.), *Studies on the acquisition of Portuguese*, 115–128. Lisboa: APL/Colibri.
- Vieira, Sónia. 2011. *A test for sentence development in European Portuguese (STSD-PT)*. Lisbon: University of Lisbon dissertation.
- Villata, Sandra, Luigi Rizzi & Julie Franck. 2016. Intervention effects and relativized minimality: New experimental evidence from graded judgments. *Lingua* 179, 76–96.
- Wexler, Kenneth. 2004. Theory of phasal development: Perfection in child grammar. *MIT Working Papers in Linguistics* 48, 159–209.
- Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. <https://ggplot2.tidyverse.org>.