



OPEN ACCESS

EDITED BY Kleanthes K. Grohmann. University of Cyprus, Cyprus

REVIEWED BY Natalia Mitrofanova, UiT The Arctic University of Norway, Norway Jeanine Treffers-Daller, University of Reading, United Kingdom

*CORRESPONDENCE Jacopo Torregrossa

RECEIVED 12 September 2024 ACCEPTED 15 October 2025 PUBLISHED 11 November 2025

CITATION

Torregrossa J, Flores C and Rinke E (2025) The relevance of instruction, language exposure and age for heritage children's development of complex morphosyntax: triangulating data from narratives and cloze-tests. Front. Lang. Sci. 4:1495607. doi: 10.3389/flang.2025.1495607

© 2025 Torregrossa, Flores and Rinke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these

The relevance of instruction, language exposure and age for heritage children's development of complex morphosyntax: triangulating data from narratives and cloze-tests

Jacopo Torregrossa^{1,2*}, Cristina Flores³ and Esther Rinke¹

¹Goethe-Universität, Frankfurt am Main, Germany, ²IDeA Center for Research on Individual Development and Adaptive Education of Children at Risk, Frankfurt am Main, Germany, ³Centro de Estudos Humanísticos, Universidade do Minho, Braga, Portugal

For children speaking a heritage language, the onset of schooling may induce a shift in dominance of language exposure from the heritage language to the societal language. This shift may affect the acquisition of morphosyntactic structures in the heritage language, especially of those that, due to their morphosyntactic complexity, tend to be acquired in school age even by monolingual children. The present study investigates the morphosyntactic abilities of 59 children aged 7 to 16 living in Switzerland, speaking European Portuguese as heritage language and German as societal language. Children's morphosyntactic abilities were measured using a written cloze-test and a narration task. By triangulating the results from both tasks, we examined which factors influenced morphosyntactic development in the heritage language, focusing on the role of task type, age, amount of heritage language instruction and variety of heritage language exposure at home, and their interaction with the complexity of the target structures. The results revealed a clear effect of complexity, with more complex structures being produced less accurately, and a task effect, with narratives yielding fewer accurate productions than clozetests. Additionally, we found a general effect of age, particularly for less complex structures. Crucially, the amount of years attending heritage language classes emerged as a key factor in the mastery of the most complex structures, whereas the variety of exposure to the heritage language at home showed no significant effect. The results are discussed in terms of the importance of exposure to formal registers of European Portuguese, provided by the heritage language courses in Switzerland, for the consolidation of more complex morphosyntactic abilities.

heritage language development, instruction, language exposure, age effects, morphosyntactic complexity, cloze-test, written narratives

1 Introduction

Across the studies on the acquisition of a heritage language (HL, henceforth), there has been an increasing interest in examining HL children's language development in the school-age period (see Montrul, 2018 for discussion). In particular, investigating the variables that affect the development of the HL during this period contributes

to explain the variation that is usually observed among adult speakers of a HL.

A focus on the school-age period may be crucial for several aspects. First, in this period (independently of schooling), children experience significant cognitive-maturational changes (see, e.g., Anderson, 2002 on the development of executive functions). Some of these changes may be relevant for the processing and acquisition of linguistic structures that are complex from a computational point of view, such as object relative clauses or certain types of object questions (Tsimpli, 2014 for a review). Second, across the literature on monolingual language acquisition, it has been often neglected that certain structures of a language are acquired—or at least, consolidated—via literacy exposure at school (Guasti and Cardinaletti, 2003). In other words, exposure to formal registers and oral and written decontextualized language use seem to be necessary for certain aspects of a language to be consolidated. For HL speakers, the relevance of the schooling period for HL development is twofold. The onset of schooling in the dominant societal language (SL) usually corresponds to a shift in language exposure: while exposure to the HL diminishes, input to the SL increases. This is due to the impact of schooling and the growing number of socialization contexts in which the SL is used (e.g., leisure activities, see Caloi and Torregrossa, 2021 for discussion). In addition, schooling is often restricted to the SL only. HL speakers therefore tend to have reduced exposure to formal registers in their HL, which are more prevalent in the classroom context through the registers used by teachers and in written texts. This may have an impact on their language acquisition processes and outcomes (Bongartz and Torregrossa, 2020; Rinke and Flores, 2014).

The present study seeks to deepen our understanding of language development in HL children of school age. It capitalizes on our previous work related to the understanding of which structures are particularly difficult in the monolingual and bilingual acquisition of European Portuguese (EP; Torregrossa et al., 2023b; Rinke et al., 2024). We aim to explore whether the observed difficulties with specific linguistic structures stem from task effects or the inherent complexity of these structures. To achieve this, we triangulate data from two sources: cloze-tests, used in prior research to identify difficult structures, and narratives, which form the primary dataset for this study. Our analysis will focus on two key questions: (1) Do children correctly use complex morphosyntactic structures in cloze-tests although these structures are absent from their narratives, where they have greater freedom in selecting preferred forms? (2) Conversely, are structures that are less accurately used in the cloze-tests nevertheless produced correctly in the narratives? By comparing evidence from these two task types, we aim to provide a more comprehensive picture of HL children's morphosyntactic knowledge than what could be gained by examining a single task in isolation. Additionally, we investigate how various factors—such as age, amount of home and school exposure to the HL-influence the production of complex morphosyntactic forms among school-aged HL children. Sensitivity to these factors may differ depending on the level of complexity of the structure.

2 Background

2.1 The acquisition of complex syntax in a heritage language

The acquisition of certain morphosyntactic structures may be challenging for HL speakers (Montrul and Polinsky, 2021; Polinsky et al., 2024 for a review). We consider here studies on HL children from very young age (4 years) to adolescence.

Recent studies using sentence repetition tasks identified morphosyntactic structures that are particularly difficult in the acquisition of a HL (Marinis and Armon-Lotem, 2015). Sentence repetition tasks serve to assess the acquisition of structures of increasing levels of complexity (e.g., simple sentences, coordinated clauses, embedded clauses, clauses involving movement, and clauses involving both movement and embedding). For example, Komeili et al. (2020) tested 20 typically developing children speaking Farsi as a HL (with English as SL) ranging in age between 6 and 11 years. They showed that the children were very accurate in the production of 'simple' structures such as short passive sentences and who-object clauses but less accurate in the production of more complex clauses such as long actional passives and different types of subject and object relative clauses. The study by Torregrossa et al. (2024) included 32 children speaking Italian as HL and German as SL ranging in age between 7 and 12 years. The authors found that these children tended to be less accurate with structures involving movement (e.g., which-object clauses and topicalizations) and structures involving both embedding and movement (e.g., relative clauses or indirect interrogative clauses) than clauses only involving embedding (e.g., complement clauses and adverbial clauses)—see also Makrodimitris and Schulz (2021) for similar results. Correia et al. (2024) tested 25 children speaking EP as HL and German as SL between 6 and 10 years with a sentence repetition task in Portuguese and found that accusative clitics (in enclisis and proclisis position), followed by conditional clauses, which-object clauses and long actional passive clauses were particularly challenging for these children.

The results reported in the abovementioned studies are in line with previous research showing that syntax acquisition is modulated by the complexity of the morphosyntactic structure to be acquired [e.g., Chondrogianni and Marinis (2011) on early L2 learners of English; Tsimpli (2014) on monolingual and bilingual acquisition]. Complex morphosyntactic structures may differ from one language to the other: what is complex in one language is not necessarily complex in another. For example, pronominal object clitics in Portuguese are acquired later than in Greek or Italian (Varlokosta et al., 2016), which suggests that these structures are associated with different degrees of complexity across the three languages (Section 3.2.3). In Section 3.2.3, we show which morphosyntactic structures are complex in EP relying on a definition of complexity that is informed by linguistic criteria.

An additional factor that should be considered in the assessment of children's acquisition of morphosyntax is the complexity of the task used. For example, Chondrogianni and Marinis (2011) administered the syntax subtest of the DELV-Norm Referenced task to child L2 learners of English and English

monolingual children ranging in age between 7 and 10 years. The L2 learners of English had an age of onset to English between 2 and 5 years. The task assessed comprehension and production of syntactic structures such as wh-questions and passives. The L2 learners were less accurate in both tasks than their monolingual peers. In contrast, Paradis and Kirova (2014) assessed the production of complex syntax among 5-year-old L2 learners of English using a story-telling task and found no significant difference between these children and their English monolingual peers (see Cahill et al., 2020; Paradis et al., 2017 for similar results). Korade et al. (2024) interpreted these results as showing that unstructured, natural and ecological tasks like storytelling may be cognitively less demanding for bilinguals than highly structured and decontextualized tasks like those related to the comprehension and production of specific target structures, as the ones used in Chondrogianni and Marinis (2011) or Marinis and Armon-Lotem (2015) [see Torregrossa et al. (2024) on ecological criteria in language assessment]. However, contrasting with the authors' interpretation of the results of previous studies, Korade et al. (2024) found that French-English bilingual children between 4 and 7 years performed better than their English monolingual peers in a structured task (i.e., a standardized complex-syntax production task based on picture stimuli) and that the two groups of children did not differ from each other in an unstructured task (a story-telling task).

In the current study, we administered a written cloze-test and a written story-retelling task to children speaking EP as HL.

2.2 Factors affecting the acquisition of morphosyntax by HL children

HL acquisition occurs under specific exposure conditions. Children acquiring a HL are exposed to it from birth within their families. Depending on the language constellation of the nuclear families and the native languages spoken by the parents, the HL may be in daily contact with other languages, particularly the dominant SL, from birth or from the moment the child starts to interact in the SL environment, for example, when entering kindergarten. A typical outcome of growing up with more than one language is that exposure to the HL decreases with increasing socialization in the dominant SL, especially when it is the main language of schooling (Section 1). As recent research on HL development has extensively shown, the particular input conditions modulating HL acquisition lead to considerable variation in the acquisition process and outcomes of the HL (Caloi and Torregrossa, 2021; Correia et al., 2024; Daskalaki et al., 2019; Flores et al., 2017; Gordon and Meir, 2024; Paradis, 2011, 2023).

Individual differences between HSs may be influenced by linguistic factors, such as the similarity of the contact languages or the complexity of the target linguistic structures (Section 2.4), and extralinguistic factors, such as the time spent with speakers of the HL, the frequency of using it in daily life and the diversity of contexts where the HL is employed (e.g., leisure activities) [Correia et al., 2024; Torregrossa et al., 2021; see De Cat et al. (2023) for an overview of variables affecting multilingual children's

language experience]. The age at which children start learning the SL may also be crucial, whereby a later age of acquisition of the SL seems to be associated with better language abilities in the HL [Armon-Lotem et al., 2021; Gagarina and Klassert, 2018; but see Armon-Lotem et al. (2011) for evidence against this generalization]. Furthermore, literacy or formal instruction in the HL affects its acquisition, as it provides children with opportunities to engage with different types of registers and various sources of HL input, thereby increasing the amount and quality of HL exposure (Armstrong, 2024; Bello-Uriarte, 2022; Rinke et al., 2019; Montrul and Armstrong, 2024 for a review).

Although much is already known about the external factors modulating HL development, research is still in its early stages of understanding how different linguistic properties and language modules are affected by varying conditions of language exposure. We know that variation in language exposure, particularly reduced contact with the HL, impacts on HL development across various domains, from phonology to vocabulary and morphosyntax (Flores et al., 2017; Hoff et al., 2012; Kupisch et al., 2021; Paradis, 2011; Rodina et al., 2020; Torregrossa et al., 2021). In general, the lower the amount of language exposure in the HL, the higher the degree of foreign accent, the smaller children's lexical repertoire, and the slower the acquisition of morphosyntactic structures in the HL.

However, these general tendencies may be modulated by other factors that may differ from one domain to the other. For example, in the lexical domain, the frequency of the lexical item and the number of cognates (form-similar translation equivalents) shared by a bilingual's pair of languages strongly interacts with language exposure variables in vocabulary growth (Garcia-Castro et al., 2023). In morphosyntax, the degree and type of language exposure interacts with the timing of acquisition and the degree of complexity of the structures to be acquired (Flores et al., 2017; Listanti and Torregrossa, 2023; Rinke et al., 2024; Schulz and Grimm, 2019). For example, Correia et al. (2024) showed that the effect of richness of input in the HL on the acquisition of morphosyntax by children speaking EP as HL (with German as SL) is particularly visible in association with syntactic structures that are considered to be more complex and are late-acquired in monolingual EP (e.g., clitics or the subjunctive). Torregrossa et al. (2023a) showed that the acquisition of structures of different degrees of complexity in a HL (as measured by using a sentence repetition task, see above) is sensitive to language exposure variables related to the time in which the target structures are known to be acquired in monolingual language acquisition. This suggests that not only language exposure plays a role, but also language exposure in specific critical periods. Along similar lines, Torregrossa et al. (2023b) showed that amount of exposure to more formal language registers in HL classes can be beneficial for stabilizing knowledge of linguistic structures that are more complex and late acquired in the acquisition of EP as HL [see Rinke and Flores (2014) for a similar consideration]. This is also relevant in the case of the acquisition of cohesive ties (e.g., referring expressions and discourse connectors) in the production of oral narratives (Bongartz and Torregrossa, 2020).

Another important consideration concerns which measure of language exposure best predicts HL children's ability in complex syntax. In this study, we focused on the variety of HL exposure

at home, measured by the number of different individuals who speak the HL within the household. This measure reflects not only the quantity, but also the quality of language exposure, since it captures the diversity of linguistic input available to the child. In our previous work, we showed that for HL speakers living in large and linguistically active communities—such as the Portuguese-speaking community in Switzerland examined here—the variety of HL exposure at home has an impact on the acquisition of complex syntax [Torregrossa et al., 2023b; see also Correia et al. (2024) for the context in Germany].

The positive effects of language exposure in the classroom context on language acquisition and processing have been documented in several studies on both child and adult monolinguals, revealing a highly variegated picture in terms of linguistic domains that are involved. For instance, literacy exposure seems to enhance comprehenders' ability to build syntactic predictions in spoken and written language processing (Favier et al., 2021). Montrul and Armstrong (2024) argue for a similar effect of literacy exposure in HL acquisition, suggesting that higher levels of instruction in the HL strengthen morphosyntactic representations, leading to more in-depth processing of information. Coming back to studies on monolinguals, schooling also seems to affect language production abilities positively, whereby speakers with higher degrees of instruction exhibit richer vocabulary knowledge (Dabrowska, 2018) and tend to use more complex structures [e.g., passive relative clauses in the study by Montag and MacDonald (2015)]. This suggests that rather than being limited to specific linguistic domains, amount of instruction may have a spreading effect across several domains.

In the present study, we administered the children two tests tapping into different linguistic abilities (including productive and receptive abilities). If the effect of amount of HL instruction is visible across different domains, it should affect their performance in both tests. In the next section, we will describe which requirements are associated with performance in each test.

2.3 Cloze-tests and narratives as instruments to assess the acquisition of syntax in a HL

In this study, we used written cloze-tests and narratives to assess HL children's morphosyntactic knowledge. Cloze-tests consist of texts with blanks, which are generated by deleting parts of or whole words (Alderson, 1979; Brown, 1980). For this study, we adopted a rational method for deletion, whereby we deleted specific morphemes or target words with lexical or grammatical meaning (e.g., nouns, prepositions, nominal, verbal markers, complementizers). This method allowed us to assess children's knowledge in specific morphosyntactic domains, focusing on the ones that previous studies have shown to be particularly vulnerable in the acquisition of a HL.

Cloze-tests are considered integrative proficiency assessment instruments because they tap into both receptive and productive lexical and morphosyntactic knowledge. To be able to fill a gap, participants need to identify the target word/morpheme based on

contextual cues and produce it accurately (Drackert and Timukova, 2020; Tremblay, 2011; Tremblay and Garrison, 2008). Therefore, cloze-tests require text comprehension abilities, morphosyntactic predictive processing, and production skills, as participants must select the appropriate morphosyntactic structure. Recent studies demonstrate that cloze-tests are reliable instruments for assessing HL speakers' morphosyntactic knowledge (Luchkina et al., 2021; Torregrossa et al., 2023b), particularly when orthographic errors are disregarded. Orthographic knowledge may be vulnerable in some HL speakers due to reduced literacy exposure (see Section 2.2). Using the terminology of Korade et al. (2024)—as discussed in Section 2.1—cloze-tests are considered 'structured' measures of morphosyntactic knowledge, as they "require specific responses within a decontextualized testing context" (Korade et al., 2024, p. 455). Consequently, they effectively enhance focus on form, particularly with respect to the target structures.

Narratives have also been widely used in studies on bilingual children's language acquisition to assess various linguistic abilities, including story structure, morphosyntactic complexity and accuracy, lexical richness, and discourse features related more generally to coherence and cohesion or, more specifically, to the distribution of null and overt subjects (Bongartz and Torregrossa, 2020; Flores and Rinke, 2025; Gagarina et al., 2015). Narrative production, in particular, requires participants to organize coherent information at a global level (Berman, 2008; Bongartz and Torregrossa, 2020; Gagarina, 2016). At the local level, it involves the use of cohesive devices reflecting the global narrative structure, such as appropriate referring expressions or complex syntax related to clause linking (Frizelle et al., 2018; Torregrossa et al., 2021; Tsimpli et al., 2016).

Narratives can also be categorized as structured tasks, as they incorporate story structure elements like goals, attempts, and outcomes (Gagarina et al., 2015). However, unlike cloze-tests, they allow participants the freedom to choose their preferred morphosyntactic structures rather than constraining them to a single option. For the present study, we elicited the narratives using pictures (see Section 3.2.2) in a retelling mode. Participants were asked to first listen to the story and then write their own version. The retelling mode allowed us to control for potential variation in participants' interpretation of the pictures, and consequently, in the content of their narrative. However, previous research has demonstrated that exposure to a model story does not influence participants' lexical choices or their use of morphosyntactic structures (Andreou et al., 2020; Gagarina et al., 2015; Listanti and Torregrossa, 2023; Otwinowska et al., 2018).

Both the cloze-tests and the narratives were administered in written form in pencil-paper format to ensure comparability between the two tasks. This approach accounted for the potential influence of transcription skills (e.g., spelling and handwriting) on participants' performance in both tasks (McIntyre et al., 2024). Orthographic errors in the written narratives were disregarded to focus on participants' morphosyntactic abilities rather than their orthographic accuracy. To maintain coherence across tasks, the cloze-test was also designed as a narrative task (see Section 3.1). Notably, all participants were able to carry out both written tasks. They had been exposed to written registers in the HL through their attendance at classes of EP as HL (Section 3.1), where the tasks

were administered. The aim of these classes is to enhance children's ability to read and write in the HL and to deepen their engagement with Portuguese culture.

A final methodological consideration concerns the advantage of triangulating the results of cloze-tests and narratives. Many studies on HL speakers' morphosyntactic abilities rely on elicited production. Frequently, structures that are absent or less commonly used by HL speakers are interpreted as evidence of erosion or simplification. This tendency has contributed to the characterization of HLs as undergoing grammatical simplification, particularly in inflectional morphology and syntax [see e.g., Brehmer (2021), Kang and Yoon (2021), and discussion in Laleko and Scontras (2021), Montrul (2016)]. Such conclusions are often extended to corpus data from both adults and children. For example, Alexiadou and Rizou (2022) analyzed production data from adolescent and adult monolingual and HL speakers of Greek with a focus on relative clauses. They show that HL speakers avoid the use of a particular type of restrictive relative clause (introduced by the pronoun o opios), arguing that this avoidance signaled morphological simplification and reduction (since this type of relativization requires overt agreement). Looking at the oral narratives produced by adult speakers in Germany/Austria speaking Bosnian as HL, Jažić et al. (2023) found that these speakers globally showed high rates of case marking accuracy (97%). Similar to the abovementioned studies, the (reduced) instances of case-marking errors (e.g., nominative-accusative substitutions) were interpreted as potentially signaling "a systematic restructuring of the underlying heritage grammar" (p.12).

While avoidance patterns and variation in usage may reflect difficulties with certain structures, they do not necessarily provide direct evidence of underlying grammatical representations. Therefore, interpreting narrative data in conjunction with clozetest results—which more explicitly target knowledge of specific morphosyntactic structures—can offer a clearer and more nuanced understanding of HL speakers' morphosyntactic abilities.

3 The present study

In this study, we examined the production of written narratives and the answers in a cloze-test in EP by bilingual children speaking EP as HL and German as SL. The study had two main aims.

First, we intended to investigate whether children produce accurate morphosyntactic structures of different levels of complexity to the same extent across the two tests. For the definition of 'complex' structures, we refer to the criteria in Section 3.2.3. We predicted that the children would produce fewer complex structures in both tests, based on the observation made in Section 2.1 that the morphosyntactic complexity of a target structure affects the pace and outcome of its acquisition (both among monolinguals and bilingual children).

Furthermore, we examined whether the production of morphosyntactic structures of different levels of complexity differed based on the nature of the task (cloze-tests vs. narratives). We expected the participants to produce more target structures and be more accurate in the cloze-tests compared to the narratives, because the former are more controlled tests and

enhance focus on form. We also anticipated that the complexity of the morphosyntactic structure would modulate the effect of task type, with the advantage of cloze tests being particularly pronounced for the most complex structures. By triangulating the analysis of narratives and cloze-tests, we were in a position to understand whether the non-occurrence of a target structure in the narratives reflected lack of mastery or just a 'sample issue', whereby children did not produce the structure even if they knew it (Section 2.3).

Finally, we aimed to identify which exposure variables affected the production of morphosyntactic structures of different levels of complexity, focusing on the use of the HL at home—especially the number of family members with whom children spoke Portuguese-and the amount of instruction in the HL classes, measured through the number of years they attended the HL classes. We also examined the role of current age at testing. Based on the studies reviewed in Section 2.2, our prediction was that older children and those with greater language and literacy exposure to the HL would produce a greater amount of accurate structures across the two tasks. However, we also expected the effect of age and HL exposure on the accurate production of morphosyntactic structures to vary based on the degree of complexity of these structures. For example, amount of HL instruction should be more relevant for the acquisition of morphosyntactic structures exhibiting a higher degree of complexity, just like in monolingual acquisition, because exposure to formal (alias written) language is beneficial to consolidate these structures. In contrast, we did not anticipate any interaction between task type and amount of HL exposure, as literacy in a HL is generally associated with a broad, spreading effect across productive and receptive language abilities, as assessed by both tasks in this study (Section 2.2). Likewise, we did not advance specific hypotheses about potential interactions between task type and either the variety of HL exposure at home or children's age.

3.1 Participants

We tested 59 children with Portuguese as HL, Swiss German as SL and Standard German as school language ranging in age from 7 years 10 months to 16 years ($M_{age} = 12$ years 2 months, SD = 21 months). We recruited them from the HL classrooms offered by the Instituto Camões (Gonçalves and Vinzentin, 2021). Among the participants, 11 were simultaneous bilinguals, exposed to both Portuguese and German from birth, with a mean length of exposure to German (LoE_GE) of 11 years and 2 months (SD = 13.69 months; range: 118-161 months). The majority were successive bilinguals, first exposed to German at different ages. Of these, 24 participants began exposure between the ages of 3 and 5, with a mean LoE_GE of 8 years (SD = 22.39 months; range: 46-143 months). Another 9 participants began exposure between the ages of 6 and 7, with a mean LoE_GE of 6 years and 10 months (SD = 25.64 months; range: 40-115 months). Additionally, three participants began exposure to German later: one at 8 years (LoE_GE: 60 months), one at 9 years (LoE_GE: 47 months), and

one at 10 years (LoE_GE: 18 months)¹. Before we conducted the study, the parents provided written informed consent. The parents and the teachers reported that none of the participants had prior identified speech, hearing or visual impairment. The study was approved by the ethics committee for Social and Human Sciences of the University of Minho (reference CEICSH 016/2019).

3.2 Research instruments and their analysis

The participants were tested in Portuguese using a written narrative-retelling task and a written cloze-test. The order of administration of these two tasks was counterbalanced across participants. Their parents or guardians filled out a language background questionnaire before we conducted the study.

3.2.1 Language background questionnaire

The language background questionnaires were administered to the participants' parents or guardians. We used an adapted version of the questionnaire designed by Correia and Flores (2021), which provides information on participants' age, their language exposure to Portuguese or (Swiss) German at home, formal instruction in Portuguese in the context of HL classrooms and age of onset to Portuguese and the SL (see *Participants* above). For each child, we extracted two indexes from the questionnaires, one related to the variety of home exposure to EP (*variety-HL exposure*) and the other one to the amount of instruction in the HL in HL classrooms (*amount of HL-instruction*).

The variety-HL-exposure index was calculated based on the number of family members who regularly spoke Portuguese to the child. The questionnaire included a list of potential interlocutors: mother, father, stepmother, stepfather, up to three siblings, and up to four grandparents living in the same household. The parents or guardians could also name up to two additional individuals. For each person listed, respondents were asked to indicate how frequently they spoke Portuguese to the child. The final score reflected the number of individuals reported to speak Portuguese either "always (and never in German)", "usually (and rarely in German)" or "half in Portuguese and half in German".

The amount of HL-instruction index was calculated as the product of the number of years in which the child attended Portuguese HL classrooms, the number of hours per week of these classrooms and the average number of school weeks per year (32, for all participants). For example, if participants had been attending HL classrooms 3 hours per week for 3 years, their amount of HL-instruction index equalled 288 hours (i.e., $3 \times 3 \times 32$).

3.2.2 The narrative production task

The narrative task was administered as a written retelling task in the form of a classroom activity. The task was based on the Edmonton Narrative Norms Instrument (ENNI; Schneider et al., 2005), which encompasses two structurally equivalent stories. One story was used for the narrative production task and one for designing the cloze-test (see next section). In this way, we maintained coherence across the two tasks (Section 2.3). Both stories consisted of 13 pictures that represent a series of events. The story used for the narrative production task involved two major characters (a giraffe boy and an elephant girl) and two minor ones (an elephant lifeguard and an elephant woman).

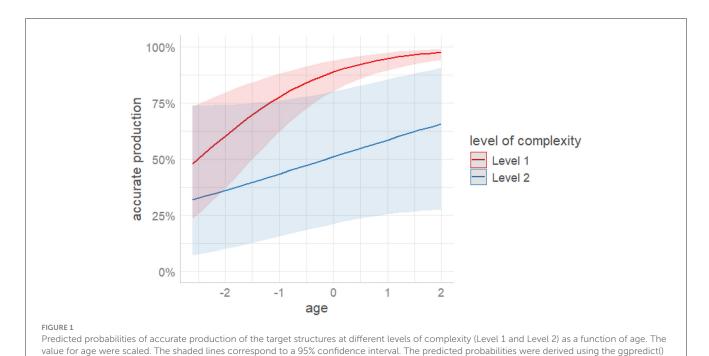
The participants were asked to listen to the story told by a prerecorded female voice while they were looking at the corresponding pictures appearing as a sequence of PowerPoint slides that were projected on the classroom's whiteboard (Torregrossa and Bongartz, 2018; Torregrossa et al., 2021 for methodology). We report the text of the model story in Supplementary material 1. Then, the participants had to retell the story writing it on a piece of paper provided to them. In this piece of paper, we reported the 13 pictures of the story, in order to help the participants remember the details of the story (see Figure 1 in Supplementary material 1 for an example of the worksheet that was provided to the participants). They were given 45 min to complete the task.

3.2.3 Structures of different degrees of complexity

For the analysis of the narratives, we focused on structures of two levels of complexity to investigate whether the participants produced them. Linguistic complexity is a challenging concept because it is theory-dependent and relational, observed indirectly through its effects. In this study, we operationalized complexity based on our previous work on monolingual and bilingual acquisition of EP. In Torregrossa et al. (2023b) we administered a cloze-test to 180 children speaking Portuguese as HL and French, German, or Italian as SLs. This test assessed the children's knowledge of different morphosyntactic structures. The analysis revealed that the structures clustered into two main groups, based on the accuracy of the responses: relatively easy structures and more difficult ones. Notably, the difficulty of these structures did not depend on the children's SLs, suggesting that the challenges were due to the linguistic features of the structures themselves. This hypothesis was further supported by the observation that EP monolingual children—who took the same cloze-test—also struggled with the same structures (Rinke et al., 2024). From this, we identified a set of 'difficult' structures, outlined in (1)-(10) below. To further understand these difficulties, we conducted a linguistically informed analysis that pinpointed the specific linguistic features contributing to their complexity (Rinke et al., 2024). In Rinke et al. (2024), the identification of difficult structures in Portuguese is framed within existing approaches to morphosyntactic complexity.

Specifically, we identified four criteria for complexity: (i) Derivational complexity (involving more layers of embedding or movement operations); (ii) Irregular patterns (i.e., exceptions requiring memory-based and item-by-item learning); (iii) Context sensitive syntactic rules; and (iv) Multiple form-function mappings

¹ Twelve parents did not provide this information in the corresponding section of the questionnaire. The three children who began exposure to German at ages 8, 9, or 10 were born and raised in Portugal and likely attended school there for 2, 3, and 4 years, respectively. Note: our questionnaire did not include questions about schooling in Portugal. Despite this variability, we included all the children in our analysis, as the variation in our sample reflects the broad range of language experiences typically observed among HSs with respect to their HL.



function in the 'ggeffects' package (Lüdecke, 2018).

(e.g., allomorphy or homophony). Several of these criteria have also been addressed in prior studies on linguistic complexity [see, for instance, Mufwene et al. (2017) for a comprehensive overview]. The 'difficult' structures identified are as follows:

- 1. Complement clauses introduced by que 'that';
- 2. Relative clauses;
- 3. Consecutive clauses introduced by the complementizer *que* 'that'.

Structures (1)–(3) involve embedding [and movement in the case of relative clauses; criterion (i)]. In addition, the multiple functions of the lexical item *que* (relative pronoun, complementizer, and interrogative element) contribute to their complexity [criterion (iv)].

- 4. Some contracted prepositions (e.g., pela 'through the');
- 5. Passive clauses.

The preposition *por* changes its form to *pelo/pela* if combined with a definite (but not an indefinite) article (= por ('through') + o/a ('the')) – criterion (iii). In addition, por as a preposition expressing a pathway (in *pela floresta* 'through the forest') involves multiple form-function mappings because of its homophony with the preposition por introducing passive agents and the lexical preposition por expressing a purpose (criterion (iv)).

6. Plural marking on irregular nouns.

Plural marking on irregular nouns can only be acquired through memory-based lexical learning [criterion (ii)].

7. Inflected infinitives.

The use of the inflected infinitive in combination with the concessive connector apesar de is complex, because apesar de exhibits multiple form-function mappings, a great variety of structures, since it combines with non-inflected such inflected and infinitival clauses. noun phrases and finite clauses, when que is [criterion (iv)].

- 8. Singular and plural dative clitics (*lhe* 'to her/to him', *lhes* 'to them'):
- 9. Singular and plural accusative clitics used in proclisis or enclisis (o 'him', a 'her', os/as 'them');
- 10. Singular and plural accusative clitics used in enclisis or mesoclisis after verbs ending in 'r' (*lo* 'him', *la* 'her', *los/las* 'them').

Structures (8)–(10) involve clitics. It is a well-known fact that third person clitic pronouns (except for reflexive clitics) are particularly complex in EP because they involve allomorphy, which represents another form of multiple form-function mappings and dependency on the phonological context [criteria (iv) and (iii), respectively]. In addition, the appropriate use of clitics generally involves discourse dependency and possibly movement [criterion (i)].

We contrasted the complex structures (1)–(10) with less complex structures, which are in general regular, unambiguous, context-independent and derivationally simple. These include:

- 11. Affixal agreement marking on adjectives;
- 12. Plural marking on regular nouns;
- 13. Superlative marking;
- 14. Modal adverbs ending in -mente;

- 15. Some contracted prepositions (e.g., \(\delta/ao\) 'to the', \(na/no\) 'in the');
- 16. Reflexive clitics (i.e., se 'herself/himself');
- 17. Temporal connectors (e.g., enquanto 'while', quando 'when');
- 18. Prepositional infinitival clauses (e.g., *para brincar* 'to play', *sem falar* 'without speaking').

For each of the structures (1)–(18), we assigned a score of 1 if they appeared in the narratives at least once and 0 if they did not.

3.2.4 Error analysis of the narratives

Before analyzing the narratives, we checked the texts produced by the children for possible grammatical errors or inaccuracies, to determine whether the target structures in the narratives were produced accurately. We found a residual number of grammatical errors across all narratives. Some of these errors did not involve the target structures. For example, some children omitted the article of avião ('airplane') writing, for instance, com avião ('with airplane') instead of com o/um avião ('with the/an airplane'). We found three instances of this error across all 59 children, two of which involved the preposition com. We also encountered three instances of errors with verbal agreement, where verbs were left uninflected (e.g., como o Girafito *princar--'how the giraffe boy play_{INF}'—instead of como o Girafito brincava—'how the giraffe boy was playing'). Other target structures that were produced incorrectly were: five incorrect productions of the periphrastic verb phrase começar a + infinitive ('start to'), with omission of a ('to'), as in começo (*a) jogar ('started play'); three incorrect uses of dative clitics with the verb ajudar ('help'), which selects accusative case; three errors in the verb form combined with an allomorphic clitic form (e.g., *ajudar-los instead of ajudálos - 'to help them'); three errors with prepositions (i.e., one omission of com 'with' in *ficou inweja ('became jealousy') instead of ficou com inveja ('became jealous'); one substitution of the contracted proposition na (=em+a, 'in(to)+the') with the preposition da = de+a, 'of the' in cair (*da) água ('fall of the water') instead of cair na água ('fall into the water'); one gender agreement error—i.e., *no piscina ('in_{MASC}-the swimming pool') instead of na piscina ('inFEM-the swimming pool')). Among the target structures, we found the following few inaccuracies: two omissions of the complementizer que, as in Elefantina wio (*que) u Girafo tina ... ('Elephantina saw (that) Giraffo had...'); one omission of the reflexive se (*sentiu mal ('felt bad'), instead of sentiu-se mal); three omissions of the complex preposition ao (=a+o, 'to-the'), as in tiro o avião (*a)o Girafo ('throw the airplane on Giraffo'), which could also reflect a phonetic-driven orthographic error, since 'ao' is often pronounced as [3]. Overall, the children produced very few grammatical errors, with even fewer occurring in the target structures. Finally, we highlight that many of the deviations reported above are also commonly found in the colloquial language use of monolingually raised children and adults (such as substituting the accusative with the dative clitic with the verb ajudar 'to help').

As mentioned above, our analysis did not consider orthographic errors. This was done for two reasons. First, our focus was on children's morphosyntactic abilities rather than orthographic skills. Second, we did not consider these errors in the cloze-test either (see below), thus ensuring consistency in the analysis of both tests [see, however, Flores et al. (2022b), for a more systematic analysis of orthographic errors in these narratives].

3.2.5 The cloze-test

The participants considered in the present study correspond to a subset of the ones considered in Torregrossa et al. (2023b). In particular, we focus here on the group of Portuguese-German bilingual children [n=60 in Torregrossa et al. (2023b); n=59 here because one child did not write the narrative]. As shown in Torregrossa et al. (2023b), all these children performed a written cloze-test, in order to assess their proficiency in Portuguese concerning different linguistic domains, i.e., nominal, verbal, prepositional and sentential domain.

We designed the cloze-test based on the ENNI story that was not used for the narrative elicitation task (see section above). This story involves two main characters (a female dog and a rabbit) and two minor ones (an older rabbit who sells balloons and a rabbit mother). We followed a rational procedure, which allowed us to tap into participants' knowledge of specific morphosyntactic structures (Section 2.3). In particular, we assessed participants' knowledge of both functional and lexical words. For functional words, the whole word was gapped or only the first letter was provided. For lexical words, the first half of the word was provided (see Supplementary material 2 for an example of the test). The test included 40 items [see Torregrossa et al. (2023b) for further methodological details, including the item analysis].

The children were asked to complete the entire cloze-test. However, for the present study, we considered only the items corresponding to the structures that we examined in the narratives (see section above). In particular, we excluded the structures of the cloze-test that were most unlikely to occur in the narratives, because, for instance, the story did not provide any context for these structures to be produced (e.g., lexical words such as repara 'observes' or sol 'sun' or prepositions occurring in a phrasal verb such as com 'with' occurring in ir ter com, meaning 'go to see'; see Supplementary material 1). This choice was related to the research question concerning whether the nature of the task (cloze-tests vs. narratives) affected the production of morphosyntactic structures of different degrees of complexity. We aimed to compare clozetests and narratives, avoiding confounding effects due to sampling inconsistencies across tasks. Overall, we considered in the clozetest analysis only the 18 morphosyntactic structures that were also analyzed in the narratives.

We coded the participants' responses as accurate (1) or inaccurate (0, meaning incorrect or missing), according to the coding criteria introduced in Torregrossa et al. (2023b). Accurate responses were either expected responses or responses that were not expected but still accurate in the given context, based on the intuitions of two native speakers (one of the authors of this study and an independent coder). For example, we expected participants to complete item [20] by writing *ao* 'to the' (see Supplementary material 2). However, the response *do* 'of the' was considered as accurate, too. Responses containing spelling errors were coded as correct, even if a participant did not consider

the number of gaps (e.g., *lindisimo instead of lindissimo 'very beautiful'), based on the considerations made in Section 2.3.

3.3 Statistical analysis

We used R (R Core Team, 2023) and lme4 (Bates et al., 2015) to fit a generalized linear mixed-effects model with the accurate production of the target structures (0 = inaccurate vs. 1 = accurate) as dependent variable. It is important to note that production accuracy reflects slightly different phenomena across the two tasks: in the cloze-test, it refers to the provision of the expected target item; in the narrative task, it refers to the spontaneous production of the relevant structure. Nevertheless, given our research interest in task effects, we use production accuracy as an umbrella term for the dependent variable across both tasks 2. We included in the model the two-way interactions between the level of complexity of the target structure (Level 1 vs. Level 2, based on the considerations made in Section 3.2.3 above) and each of the following predictors: type of task (cloze-test vs. narrative), variety of HL exposure at home, amount of HL instruction, and age. We fit the model with by-participant and by-structure random slopes for type of test. The variables age, variety of HL exposure at home and amount of HL instruction were scaled with the function scale() in R. The variable related to the level of complexity of the target structures was contrast coded as -0.5 for Level 1 and +0.5 for Level 2. The variable related to the type of task was contrast coded as -0.5 for cloze-test and +0.5 for narratives³.

4 Results

4.1 Language background questionnaire

The values of the *variety-HL-exposure* index ranged between 1 and 5 (M=2.58, SD = 0.90). The values of the *amount of HL-instruction* index ranged between 32 and 576 (h) (M=226.91, SD = 134.75). The two variables did not correlate with each other [$r_{(52)}=-0.16$, p=0.25]. Age correlated with the *amount of HL-instruction* index [$r_{(52)}=0.46$, p<0.001], but not with the *variety-HL-exposure* index [$r_{(52)}=0.14$, p=0.31].

TABLE 1 List of all target structures sorted out based on their degree of morphosyntactic complexity (Level 1 vs. Level 2).

Structures	Accurate production in the narratives (n. participants)	Accurate production in the cloze-test (n. participants)				
Level 1 of morphosyntactic complexity						
Agreement marking (adjectives)	50	50				
Plural marking (regular nouns)	25	57				
Superlative marking	3	36				
Modal adverbs	7	46				
Contracted prepositions (e.g., à/ao/na/no)	51	39				
Reflexive clitics	37	46				
Temporal connectors	14	42				
Prepositional infinitival clauses	18	52				
Level 2 of morphosyn	tactic complexity					
Complement clauses	41	40				
Relative clauses	28	20				
Consecutive clauses	11	36				
Contracted prepositions (e.g., <i>pela</i>)	1	25				
Plural marking (irregular nouns)	1	38				
Passive structures	0	25				
Inflected infinitives (in concessive structures)	2	33				
Dative clitics	21	30				
Accusative clitics [o/a/os/as]	18	31				
Accusative clitic allomorphs after <r>, <s>, <z> [lo/la/los/las]</z></s></r>	7	26				

For each target structure, the table reports the number of participants who produced it accurately in the narratives and in the cloze-test, respectively.

4.2 Analysis for each target structure: accurate production in the narrative and in the cloze-test

Table 1 reports for each target structure, the number of participants who produced it accurately in the narratives and in the cloze-tests, respectively. The structures are divided based on their level of morphosyntactic complexity (Level 1 and Level 2).

The table shows that structures that did not occur in the narrative were nevertheless correctly produced in the cloze test. This is expected and suggests that the narrative-production data did not fully account for participants' morphosyntactic knowledge: The number of participants who produced the target structures tended to be lower than the number of participants who provided

² Nevertheless, in the Supplementary material we report separate analyses for the cloze test and the narrative task, examining the effects of age, variety of HL exposure at home, and HL instruction. These additional analyses account for the slight differences in how the dependent variable is defined across tasks. Notably, the results reveal the same overall tendencies as those presented in the main analysis, although some interactions do not reach significance—likely due to the reduced statistical power when the data are split (see Supplementary material for further discussion).

³ The resulting model was: m0 = glmer (accurate production ~ 1 + structure complexity * (type of task + age + variety-HL-exposure + HL-literacy-exposure) + (1+test|ID) + (1+test|structure), data = d, family = binomial, control = glmerControl(optimizer = "bobyqa"))

an accurate answer in association with these structures in the clozetests. For instance, with the structures at Level 1 of morphosyntactic complexity, the participants were generally very accurate in the cloze-test, with the number of participants answering correctly to the corresponding items ranging from 36 (i.e., 66% of participants, in association with superlative marking) to 57 (i.e., 97% of participants in association with plural marking on regular nouns). However, many participants did not use the target structures in the narratives, despite producing them accurately in the clozetest. For example, the number of participants who produced superlatives, modal adverbs, prepositional infinitival clauses and temporal connectors was very low. Crucially, based on the results related to the production of infinitival clauses and temporal connectors, it might be concluded that the participants tended to simplify structure, avoiding, for instance, subordination. However, the participants showed a relatively stable knowledge of these structures in the cloze-test. This also held for morphosyntactic structures at Level 2 of complexity. The very few or no occurrences of contracted prepositions, passive structures, inflected infinitives and some types of accusative clitics in the narratives did not reflect that many participants knew these structures in the cloze-tests (e.g., 42% of participants in association with contracted prepositions and passive structures, 56% with inflected infinitives and 64% with plural marking with irregular nouns).

However, it should be noticed that in some cases, the narrative data did reflect the data drawn from the cloze-tests. For example, many more participants produced reflexive clitics than dative or accusative third person clitics in the narratives. This reflected the observation that the participants provided more accurate answers with the former than the latter in the cloze-tests. Likewise, when considering different types of subordinate clauses involving the complementizer que, participants were by far more likely to produce complement clauses than relative or consecutive clauses, both in the cloze-tests and in the narratives. In other words, the production of different structures in the narratives reflected the hierarchy of complexity emerging from the analysis of the clozetests (in terms of number of children who provided a correct answer with each target structure). Needless to say, in the narratives, children were free to choose the structures they wanted to produce (Section 2.3).

4.3 Factors affecting participants' morphosyntactic knowledge

Table 2 reports the results of the generalized linear mixedeffects model with the accurate production of the target structures
as dependent variable and, as predictor, the interaction between
the level of complexity of the structures and each of the following
variables: type of task, age, variety-HL-exposure and amount of HLinstruction. We found a significant effect of level of complexity,
whereby more complex structures were less likely to be produced
(negative estimate). We also found a significant effect of type
of task, which indicated that the participants tended to produce
fewer structures in the narratives than in the cloze-tests (negative
estimate).

TABLE 2 Parameters of the generalized linear mixed-effects analysis concerning accurate production of the target structures (0 = inaccurate, 1 = accurate) for the interaction between complexity level of the target structures and each of the following variables: type of task, age, variety of HL exposure at home and amount of instruction in the HL.

Fixed effects	В	SE	Z	р
Intercept	-0.28	0.34	-0.81	0.42
Level of complexity (Level 2)	-2.07	0.63	-3.29	<0.001
Type of task (narrative)	-2.66	0.56	-4.75	<0.001
Amount of HL instruction	0.18	0.16	1.17	0.24
variety-HL-exposure	0.09	0.14	0.64	0.52
Age	0.57	0.17	3.32	<0.001
Level of complexity × Type of task	-0.04	1.10	-0.04	0.97
Level of complexity \times Amount of HL instruction	0.37	0.15	2.38	0.02
Level of complexity × variety-HL-exposure	-0.09	0.14	-0.67	0.50
Level of complexity × Age	-0.52	0.16	-3.19	0.001

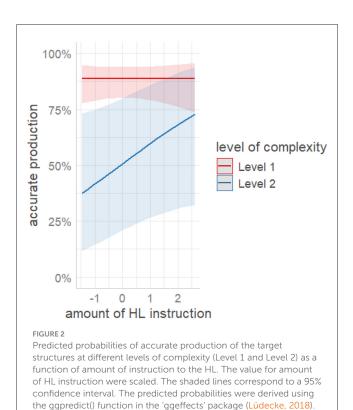
Significant effects are in bold.

We also found a significant effect of age: Older children tended to perform better than younger children across the board. Furthermore, age interacted with the level of complexity of the target structures. This indicated that less complex structures (Level 1) were more sensitive to the effect of age (negative estimate; cf. Figure 1). Using the *ggpredict()* function in the 'ggeffects' package (Lüdecke, 2018), we calculated the predicted probabilities of accurate production as a function of age and level of complexity of the target structures. The predicted probability of producing the target structures accurately increased with age by around 50% for less complex structures (Level 1) and by around 34% for more complex ones (Level 2).

Finally, we found an interaction between level of complexity of the target structures and amount of HL instruction. This indicated that the structures at Level 2 of complexity were more sensitive to the effect of amount of instruction in the HL than the structures at Level 1 (Figure 2). Marginal \mathbb{R}^2 for the model was .32, while the conditional \mathbb{R}^2 was .68. Multicollinearity was low (VIF range 1.01–2.53), indicating no concern for collinearity among the predictors, even considering the correlation between age and amount of HL instruction.

5 Discussion

In the present study, we defined 'complex' morphosyntactic structures based on linguistic criteria (Section 3.2.3). The first result emerging from the study was that, as expected, children speaking EP as HL produced fewer complex structures than simpler ones. Moreover, they produced more accurate instances of target structures in the cloze tests than in the narratives (see Tables 1, 2). This outcome likely reflects the increased linguistic demands involved in narrative production, as cloze-tests are more controlled and foster a greater focus on form. Contrary to our expectations, however, we found no evidence that the complexity



of the morphosyntactic structures modulated the effect of task type. In other words, children consistently produced fewer target structures in the narratives than in the cloze-tests, regardless of the structural complexity. Importantly, these findings suggest that the non-occurrence of morphosyntactic structures in the narratives should not be interpreted as a lack of grammatical knowledge or simplified mental grammar, but rather as a sampling issue. We return to this point in more detail below.

Against the backdrop of these findings, the relevant question is which background variables among age, variety of HL exposure at home and amount of instruction in the HL favor the acquisition of difficult structures. Our results showed that children's current age played a significant role: Older children tended to perform better than younger children with both 'easy' and 'complex' structures.

Crucially, the observation that morphosyntactic structures emerge over time suggests that the children in this study did not experience arrested development, but rather their language development was still ongoing. This contrasts with other studies which did not find an effect of age on the acquisition of morphosyntax in a HL and is presumably related to the characteristics of the community considered in this contribution. The Portuguese community in Switzerland is numerous, EP being the third most spoken language in this country (Flores et al., 2022a; Gonçalves and Vinzentin, 2021). Therefore, Portuguese-descendant children have several opportunities to speak the HL both at home and outside their home (see Rodina et al., 2020 on the positive effects of the size of the immigrant community on HL acquisition). In addition, many parents indicated in

the questionnaires that they often spend holidays in Portugal⁴. Therefore, the children considered here may have had a greater and more continuous exposure to the HL than other children in other communities with different characteristics [such as the Spanish or the Russian communities in the United States, see e.g., Benmamoun et al. (2013)]. Interestingly, we found no evidence that variety of language exposure at home affected children's acquisition of 'less complex' or 'more complex' structures in the HL. For the easiest structures, this may be because their acquisition is typically not sensitive to variables related to language exposure (Tsimpli, 2014). For the most difficult structures, one possible explanation is that, although the children showed some variability in the index of variety of language exposure at home (Section 3.1), they may have all reached a threshold sufficient for the successful acquisition of these structures. Alternatively, stabilizing knowledge of complex structures may depend more on the richness of exposure outside the home, particularly through HL courses or related factors than on exposure to the HL at home. We will revisit the role of amount of instruction below. Overall, the results of our study align with the idea that the process (and possibly the outcome) of HL acquisition mirrors that observed in monolingual children, provided the children are continuously exposed to the HL throughout their lives (Caloi and Torregrossa, 2021; Flores et al., 2017).

We also observed that age was not sufficient to explain the acquisition of the most difficult structures. Specifically, the effect of age was most evident with the easiest structures. For example, accuracy with the easiest structures in the cloze tests was nearly at ceiling as age increased (Figure 1). On the contrary, the most difficult structures (at Level 2 of complexity) seemed to be more sensitive to variables related to the variety of input sources, in particular the amount of hours of attendance of HL classrooms, which had a positive effect on accuracy in the production of the most complex structures, but not of the easiest ones (Figure 2). These findings indicate that age-related development alone does not suffice to consolidate the acquisition of more complex morphosyntactic structures. Rather, amount of HL instruction emerges as a key factor in supporting their stabilization.

⁴ Our background questionnaire included two questions in this regard: (i) How frequently does the child spend holidays in Portugal? (Options: never = 1; once every two years = 2; once a year = 3; twice a year = 4; more than twice a year = 5); (ii) How frequently does the child communicate in Portuguese during these holidays? (Options: never = 1; rarely = 2; sometimes = 3; frequently = 4; always = 5). We then multiplied the scores obtained for each answer. The participants received a score ranging from 10 (i.e., two children spending holidays in Portugal once a year and always speaking Portuguese there) to 25 (i.e., four children spending holidays in Portugal more than twice a year and always speaking Portuguese there). The majority of the children (n=28) obtained a score of 20 (twice a year, always speaking Portuguese). The mean score was 18.23 (SD = 3.57). Given the relatively homogenous nature of the sample in this regard, we decided not to include this variable in the model, especially since it was challenging to hypothesize how spending holidays in Portugal once or twice a year might interact with the level of complexity of the target structures. However, we acknowledge that some studies found an effect of holiday immersion on the development of a HL, particularly in samples with more variation in this respect (Chondrogianni and Daskalaki, 2023; Kubota and Rothman, 2024)

The positive effect of instruction is not surprising considering that monolingual children also fully master most of the structures included in the Level 2 of complexity at school age (Rinke et al., 2024 and Torregrossa et al., 2023b for discussion). This suggests that exposure to formal registers at school is beneficial for the consolidation of the most complex morphosyntactic structures in a language, regardless of the acquisition setting (monolingual or bilingual). We interpret this consolidation as stemming from the fact that, in formal instruction contexts, children have the opportunity to engage with registers in which such structures—particularly the more complex ones—occur more frequently. Formal registers also help counterbalance the variation characteristic of colloquial registers. In addition, formal instruction promotes attention to form through the development of reading, writing and grammar skills, which plays a crucial role in consolidating the most complex morphosyntactic structures (Polinsky and Kagan, 2007).

We would like to conclude with some methodological remarks. Most of the observations reported in this study were based on the triangulation of cloze-test and narrative data. Relying solely on narrative data could have led to misleading generalizations about HL children's morphosyntactic knowledge. For example, the limited production of certain Level 1 structures in the narratives (e.g., infinitival and adverbial clauses; see Table 1) might have suggested a lack of mastery. However, the cloze-test results revealed that children were in fact capable of producing these structures accurately when prompted in a more controlled context, with some exceptions observed for relative clauses at Level 2 and contracted prepositions at Level 1. Furthermore, we noticed that the children found the most complex structures challenging across both tasks (Table 1). In many cases (e.g., clitics and complementizers), the number of children who produced these structures in the narratives reflected their degree of complexity, as determined by performance in the cloze-test. Overall, our findings suggest a nuanced picture: in order to draw fine-grained conclusions about HL children's morphosyntactic knowledge, studies should triangulate (if possible) different methodologies and tasks.

As a limitation of the study, it should be noted that the sample size was relatively small for the analysis conducted, which involved multiple two-way interactions. This limitation stems from the challenge of assembling a representative group of participants for the population under investigation. Therefore, this study should be regarded as exploratory in nature.

6 Conclusion

In summary, this study offers valuable insights into the acquisition of complex morphosyntactic structures by HL children, particularly those acquiring EP in Switzerland. Our findings challenge the assumption that HL development tends to freeze during school age due to increased exposure to and schooling in the SL. Older HL children, who receive formal instruction in their HL through dedicated courses, generally outperform younger HL children in both easy and complex structures, indicating that morphosyntactic knowledge increases with increasing age. However, the acquisition of the most difficult structures appears to be more strongly influenced

by the type of exposure, especially to formal registers in HL classes, than by variety of HL exposure at home or age alone. This indicates that exposure to diverse HL input sources, and especially formal educational settings such as the Portuguese HL courses in Switzerland [see Flores et al. (2022a); Gonçalves and Vinzentin (2021) for a characterization of these courses], is crucial for the consolidation of advanced morphosyntactic knowledge. Ultimately, this study underscores the resilience of HL acquisition and emphasizes the importance of language policies that support HL education to ensure balanced minority language development [already highlighted by Valdés (1995) and many others].

Data availability statement

The data and analyses reported in this paper are available at the following OSF link: https://osf.io/5jt9n/overview?view_only=d2f8a41f64494cb79b16067018ff65f1.

Ethics statement

The studies involving humans were approved by Ethics Committee for Social and Human Sciences of the University of Minho. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

JT: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. CF: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Writing – original draft. ER: Conceptualization, Data curation, Investigation, Methodology, Resources, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was partially funded by the FCT—Fundação para a Ciência Tecnologia (Portugal), through the project https://doi.org/10.54499/UIDB/00305/2020.

Acknowledgments

We are grateful to the teachers of the Camões, Institute of Cooperation and Language, I.P., working in Switzerland, for collaborating with the data collection and to Dr. Maria de Lurdes Santos Gonçalves for coordinating their efforts. A huge thank you to the children who participated in the study and to their families.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/flang.2025. 1495607/full#supplementary-material

References

Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language. TESOL Q. 13, 219–228. doi: 10.2307/3586211

Alexiadou, A., and Rizou, V. (2022). Restrictive relative clauses in Greek heritage speakers in the US. *Bergen Lang. Linguist. Stud.* 12, 130–142. doi: 10.15845/bells.v12i2.3834

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child NeuropsycholChild Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724

Andreou, M., Torregrossa, J., and Bongartz, C. (2020). The sharing of reference strategies across two languages. The production and comprehension of referring expressions by Greek-Italian bilingual children. *Discours* 26, 1–30. doi: 10.4000/discours.10709

Armon-Lotem, S., Rose, K., and Altman, C. (2021). The development of English as a heritage language: The role of chronological age and age of onset of bilingualism. *First LanguageFirst Lang.* 41, 67–89. doi: 10.1177/0142723720929810

Armon-Lotem, S., Walters, J., and Gagarina, N. (2011). The impact of internal and external factors on linguistic performance in the home language and in L2 among Russian-Hebrew and Russian-German preschool children. *Linguist. Approaches Biling.* 1, 291–317. doi: 10.1075/lab.1.3.04arm

Armstrong, A. (2024). Literacy effects on language acquisition and sentence processing in adult L1 and school-age heritage speakers of Spanish (Unpublished doctoral dissertation). The University of Illinois at Urbana-Champaign, Illinois.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–48. doi: 10.18637/jss.v067.i01

Bello-Uriarte, A. (2022). "Effects of Instruction on specific measures of accuracy in Spanish heritage learners' writing," in *Outcomes of University Spanish Heritage Language Instruction in the United States*, ed. M. Bowles (Washington, DC: Georgetown University Press), 61–80.

Benmamoun, E., Montrul, S., and Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theor. Linguist.* 39, 129–181. doi: 10.1515/tl-2013-0009

Berman, R. (2008). The psycholinguistics of developing text construction. J. Child Lang. 35,735-771. doi: 10.1017/S0305000908008787

Bongartz, C., and Torregrossa, J. (2020). The effects of balanced biliteracy on Greek–German bilingual children's secondary discourse ability. *J. Biling. Educ. Biling.* 23, 948–963. doi: 10.1080/13670050.2017.1355888

Brehmer, B. (2021). "Slavic heritage languages around the globe," in *The Cambridge Handbook of Heritage Languages and Linguistics*, eds. S. Montrul and M. Polinsky (Cambridge: Cambridge University Press), 13–44.

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *Modern Lang. J.* 64, 311–317. doi: 10.1111/j.1540-4781.1980.tb05198.x

Cahill, P., Cleave, P., Asp, E., Squires, B., and Kay-Raining Bird, E. (2020). Measuring the complex syntax of school-aged children in language sample analysis: a known-groups validation study. *Int. J. Lang. Commun. Disord.* 55, 765–776. doi: 10.1111/1460-6984.12562

Caloi, I., and Torregrossa, J. (2021). Home and school language practices and their effects on heritage language acquisition: a view from heritage Italians in Germany. Languages 6:50. doi: 10.3390/languages6010050

Chondrogianni, V., and Daskalaki, E. (2023). Heritage language use in the country of residence matters for language maintenance, but short visits to the homeland can boost heritage language outcomes. *Front. Lang. Sci.* 2:1230408. doi: 10.3389/flang.2023.1230408

Chondrogianni, V., and Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Ling. Approach. Biling.* 1, 318–345. doi: 10.1075/lab.1.3.05cho

Correia, L., and Flores, C. (2021). Questionário sociolinguístico parental para famílias emigrantes bilingues (QuesFEB): uma ferramenta de recolha de dados sociolinguísticos de crianças falantes de herança. *Linguística. Revista de Estudos Linguísticos da Universidade do Porto* 16, 75–102. doi: 10.21747/16466195/ling16a3

Correia, L., Lobo, M., and Flores, C. (2024). A sentence repetition task for European Portuguese: results from a study with monolingual and Portuguese-German bilingual children. *Lang. Acquis.* 32, 453–483. doi: 10.1080/10489223.2024.2346586

Dabrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. Cognition 178, 222–235. doi: 10.1016/j.cognition.2018.05.018

Daskalaki, E., Chondrogianni, V., Blom, E., Argyri, F., and Paradis, J. (2019). Input effects across domains: the case of Greek subjects in child heritage language. *Second Lang. Res.* 35, 421–445. doi: 10.1177/0267658318787231

De Cat, C., Kašćelan, D., Prévost, P., Serratrice, L., Tuller, L., and Unsworth, S. (2023). How to quantify bilingual experience? Findings from a Delphi consensus survey. *Bilingual. Lang. Cogn.* 26, 112–124. doi: 10.1017/S1366728922000359

Drackert, A., and Timukova, A. (2020). What does the analysis of C-test gaps tell us about the construct of a C-test? A comparison of foreign and heritage language learners' performance. *Lang. Test.* 37, 107–132. doi: 10.1177/0265532219861042

Favier, S., Meyer, A. S., and Huettig, F. (2021). Literacy can enhance syntactic prediction in spoken language processing. *J. Exp. Psychol.* 150, 2167–2174. doi: 10.1037/xge0001042

Flores, C., Gonçalves, M. L., Rinke, E., and Torregrossa, J. (2022a). Perspetivas múltiplas sobre a competência bilingue de crianças lusodescendentes residentes na Suíça: a investigação linguística em diálogo com a didática. Revista Portuguesa De Educação 35, 102–131. doi: 10.21814/ppe.24205

Flores, C., and Rinke, E. (2025). Overt pronouns are challenging: subject realization in written narratives of Portuguese-French bilingual school-age children. *Romanica Cracoviensia* 24, 383–398. doi: 10.4467/20843917RC.24.029.20984

Flores, C., Rinke, E., Torregrossa, J., and Weingärtner, D. (2022b). Language separation and stable syntactic knowledge: Verbs and verb phrases in bilingual children's narratives. *Journal of Portuguese Linguistics*, 5, 1–30. doi: 10.16995/jpl.8043

Flores, C., Santos, A. L., Jesus, A., and Marques, R. (2017). Age and input effects in the acquisition of mood in Heritage Portuguese. *J. Child Lang.* 44, 795–828. doi: 10.1017/S0305000916000222

Frizelle, P., Thompson, P. A., McDonald, D., and Bishop, D. V. M. (2018). Growth in syntactic complexity between four years and adulthood: evidence from a narrative task. *J. Child Lang.* 45, 1174–1197. doi: 10.1017/S0305000918000144

Gagarina, N. (2016). Narratives of Russian–German preschool and primary school bilinguals: Rasskaz and Erzaehlung. *Appl. Psychol.* 37, 91–122. doi: 10.1017/S0142716415000430

Gagarina, N., and Klassert, A. (2018). Input dominance and development of home language in Russian-German bilinguals. *Front. Commun.* 3:40. doi: 10.3389/fcomm.2018.00040

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiuniene, I., et al. (2015). "Assessment of narrative abilities in bilingual children," in Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment, eds. S. Armon-Lotem, J. de Jong, and N. Meir (Bristol: Multilingual Matters).

Garcia-Castro, G., Avila-Varela, D. S., Castillejo, I., and Sebastian-Galles, N. (2023). Cognate Beginnings to Bilingual Lexical Acquisition. *PsyArXiv*, 14. doi: 10.31234/osf.io/dxsmz

Gonçalves, M., and Vinzentin, M. (2021). "Portuguese in Switzerland: investigating the representations of families and teachers," in *Portuguese as a Heritage Language in Europe: A Pluricentric Perspective*, eds. A. Souza and S. Melo-Pfeifer (Campinas: Pontes Editores), 119–150.

Gordon, S., and Meir, N. (2024). English as a heritage language: The effects of input patterns and contact with Hebrew. *Int. J. Bilingual.* 28, 353–373. doi:10.1177/13670069231155775

Guasti, M. T., and Cardinaletti, A. (2003).Relative clause formation in Romance child production. *Probus* 15, 47–89. doi: 10.1515/prbs.2003.005

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., Parra, M., et al. (2012). Dual language exposure and early bilingual development. *J. Child Lang.* 39, 1–27. doi: 10.1017/S0305000910000759

Jažić, I., Gagarina, N., and Perovic, A. (2023). Case marking is different in monolingual and heritage Bosnian in digitally elicited oral texts. *Front. Psychol.* 13:832831. doi: 10.3389/fpsyg.2022.832831

Kang, H-. S., and Yoon, J. (2021). "Heritage Languages in Japan and Korea," in *The Cambridge Handbook of Heritage Languages and Linguistics*, eds. S. Montrul and M. Polinsky (Cambridge: Cambridge University Press), 111–128.

Komeili, M., Marinis, T., Tavakoli, P., and Kazemi, Y. (2020). Sentence repetition in Farsi-English bilingual children. *J. Eur. Second Lang. Assoc.* 4, 1–12. doi: 10.22599/jesla.55

Korade, C., Nicoladis, E., and Charest, M. (2024). Monolingual and bilingual children's production of complex syntactic structures. *J. Child Lang.* 51, 454–469. doi: 10.1017/S0305000922000411

Kubota, M., and Rothman, J. (2024). Modelling individual differences in vocabulary development: a large-scale study on Japanese heritage speakers. *Child Dev.* 96, 325–340. doi: 10.1111/cdev.14168

Kupisch, T., Kolb, N., Rodina, Y., and Urek, O. (2021). Foreign accent in pre- and primary school heritage bilinguals. *Languages* 6:96. doi: 10.3390/languages6020096

Laleko, O., and Scontras, G. (2021). On the many dimensions of complexity in heritage languages. *Herit. Lang. J.* 18, 1-37. doi: 10.1163/15507076-12340009

Listanti, A., and Torregrossa, J. (2023). The production of preverbal and postverbal subjects by Italian heritage children: timing of acquisition matters. *First Lang.* 43, 431–460. doi: 10.1177/01427237231170486

Luchkina, T., Ionin, T., Lysenko, N., Stoops, A., and Suvorkina, N. (2021). Evaluating the Russian language proficiency of bilingual and second language learners of Russian. *Languages* 6:83. doi: 10.3390/languages6020083

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *J. Open Source Softw.* 3:772. doi: 10.21105/joss.00772

Makrodimitris, C., and Schulz, P. (2021). Does timing in acquisition modulate heritage children's language abilities? Evidence from the Greek LITMUS sentence repetition task. *Languages* 6:49. doi: 10.3390/languages6010049

Marinis, T., and Armon-Lotem, S. (2015). "Sentence repetition," in Assessing Multilingual Children Disentangling Bilingualism from Language Impairment, eds. S. Armon-Lotem, J. de Jong and N. Meir (Bristol: Multilingual Matters), 95–122.

McIntyre, A., Scott, A., McNeill, B., and Gillon, G. (2024). Comparing young children's oral and written story retelling: the role of ideation and transcription. *Speech Lang. Hear.* 28:2357450. doi: 10.1080/2050571X.2024.2357450

Montag, J. L., and MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8- and 12-year-old children and adults. *J. Exp. Psychol. Gener.* 144, 447–468. doi: 10.1037/xge0000054

Montrul, S. (2016). The Acquisition of Heritage Languages. Cambridge University Press.

Montrul, S. (2018). Heritage language development: connecting the dots. Int. J. Bilingual. 22, 530-546. doi: 10.1177/1367006916654368

Montrul, S., and Armstrong, A. (2024). "The Literacy Enhancement Hypothesis in bilingual language development," in Multilingual Acquisition and Learning: An Ecosystemic View to Diversity (Studies in Bilingualism; Vol. 67), ed. E. Babatsouli (John Benjamins Publishing Company), 60–89.

Montrul, S., and Polinsky, M. (2021). "Introduction: heritage languages, heritage speakers, heritage linguistics," in *The Cambridge Handbook of Heritage Languages and Linguistics*, eds. S. Montrul and M. Polinsky (Cambridge: Cambridge University Press), 1–10.

Mufwene, S. S., Coupé, C., and Pellegrino, F. (2017). "Complexity in language: a multifaceted phenomenon," in *Complexity in Language: Developmental and Evolutionary Perspectives*, eds. S. S. Mufwene, C. Coupé, and F. Pellegrino (New York, NY: CUP), 1–29.

Otwinowska, A., Mieszkowska, K., Białecka-Pikul, M., Opacki, M., and Haman, E. (2018). Retelling a model story improves the narratives of Polish-English bilingual children. *Int. J. Biling. Educ. Biling.* 23, 1083–1107. doi: 10.1080/13670050.2018.1434124

Paradis, J. (2011). Individual differences in child english second language acquisition: comparing child-internal and child-external factors. *Linguist. Approaches Biling*. 1, 213–237. doi: 10.1075/lab.1.3.01par

Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *J. Child Lang.* 50, 793–817. doi: 10.1017/S0305000922000708

Paradis, J., and Kirova, A. (2014). English second-language learners in preschool: profile effects in their English abilities and the role of home language environment. *Int. J. Behav. Dev.* 38, 342–349. doi: 10.1177/0165025414530630

Paradis, J., Rusk, B., Sorenson Duncan, T., and Govindarajan, K. (2017). Children's second language acquisition of English complex syntax: the role of age, input and cognitive factors. *Ann. Rev. Appl. Linguist.* 37, 1–20. doi: 10.1017/S0267190517000022

Polinsky, M., and Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. Lang. Linguist. Compass. 1, 368–395. doi: 10.1111/j.1749-818X.2007.00022.x

Polinsky, M., Putnam, M., and Salmons, J. (2024). "Linguistic complexity in heritage languages: an introduction," in *Formal Approaches to Complexity in Heritage Languages* (Berlin: Language Science Press), 1–13.

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Available online at: $\frac{1}{2} \frac{1}{2} \frac{$

Rinke, E., and Flores, C. (2014). Morphosyntactic knowledge of clitics by Portuguese heritage bilinguals. *Bilingual. Lang. Cogn.* 17, 681–699. doi: 10.1017/S136672891300076X

Rinke, E., Flores, C., and Santos, A. L. (2019). "Heritage languages at school: implications of linguistic research on bilingualism for heritage language teaching," in Romanische Sprachen in ihrer Vielfalt: Brückenschläge zwischen linguistischer Theoriebildung und Fremdsprachenunterricht (Romanische Sprachen und ihre Didaktik), eds. C. Gabriel, J. Grünke and S. Thiele (Stuttgart: bidem), 211–232.

Rinke, E., Flores, C., and Torregrossa, J. (2024). "How different types of complexity can account for difficult structures in bilingual and monolingual language acquisition," in *Formal Approaches to Complexity in Heritage Languages*, eds. M. Polinsky and M. Putnam (Berlin: Language Science Press), 43–71.

Rodina, Y., Kupisch, T., Meir, N., Mitrofanova, N., Urek, O., Westergaard, M., et al. (2020). Internal and external factors in heritage language acquisition: evidence from heritage Russian in Israel, Germany, Norway, Latvia and the United Kingdom. *Front. Educ.* 5:20. doi: 10.3389/feduc.2020.00020

Schneider, P., Dubé, R. V., and Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*. Edmonton, AB: University of Alberta, Faculty of Rehabilitation Medicine.

Schulz, P., and Grimm, A. (2019). The age factor revisited: timing in acquisition interacts with age of onset in bilingual acquisition. *Front. Psychol.* 9:2732. doi:10.3389/fpsyg.2018.02732

Torregrossa, J., Andreou, M., Bongartz, C., and Tsimpli, I. M. (2021). Bilingual acquisition of reference: The role of language experience, executive functions and cross-linguistic effects. *Biling. Lang. Cogn.* 24, 694–706. doi: 10.1017/S1366728920000826

Torregrossa, J., and Bongartz, C. (2018). Teasing apart the effects of dominance, transfer and processing in reference production by German-Italian bilingual adolescents. *Languages* 3:36. doi: 10.3390/languages3030036

Torregrossa, J., Caloi, I., Listanti, A., and Romano, F. (2023a). "The acquisition of syntactic structures in Italian: assessing the role of language exposure at critical periods," in *Studies in Italian as a Heritage Language*, ed. F. Romano (Berlin: De Gruyter Mouton), 155–194.

Torregrossa, J., Flores, C., and Rinke, E. (2023b). What modulates the acquisition of difficult structures in a heritage language? A study on Portuguese in contact with French, German and Italian. *Biling. Lang. Cogn.* 26, 179–192. doi:10.1017/S1366728922000438

Torregrossa, J., Listanti, A., Bongartz, C., and Marinis, T. (2024). Adding discourse to sentence repetition tasks: under which conditions does bilingual children's performance improve? *Res. Methods Appl. Linguist.* 3:100107. doi: 10.1016/j.rmal.2024.100107

Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the Gap. *Stud. Second Lang. Acquis.* 33, 339–372. doi: 10.1017/S0272263111000015

Tremblay, A., and Garrison, M. D. (2008). "Cloze tests: a tool for proficiency assessment in research on L2 French," in *Selected Proceedings of the Second Language Research Forum*, ed. MT Prior (Somerville: Cascadilla Proceedings Project), 73–88.

Tsimpli, I. (2014). Early, late or very late? Timing acquisition and bilingualism. *Ling. Approaches Biling.* 4, 283–313. doi: 10.1075/lab.4.3.01tsi

Tsimpli, I. M., Peristeri, E., and Andreou, M. (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguist*. 37, 195–216. doi: 10.1017/S0142716415000478

Valdés, G. (1995). The teaching of minority languages as academic subjects: pedagogical and theoretical challenges. *Modern Lang. J.* 79, 299–328. doi: 10.1111/j.1540-4781.1995.tb01106.x

Varlokosta, S., Belletti, A., Costa, J., Friedmann, N., Gavarró, A., Grohmann, K., et al. (2016). A cross-linguistic study of the acquisition of clitic and pronoun production. *Lang. Acquis.* 23, 1–26. doi: 10.1080/10489223.2015.1028628