



International Journal of Bilingual Education and Bilingualism

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rbeb20

Minority language exposure predicts multilingual receptive abilities in Catalan-Spanish bilingual children

Adriana Soto-Corominas

To cite this article: Adriana Soto-Corominas (02 Dec 2024): Minority language exposure predicts multilingual receptive abilities in Catalan-Spanish bilingual children, International Journal of Bilingual Education and Bilingualism, DOI: [10.1080/13670050.2024.2433139](https://doi.org/10.1080/13670050.2024.2433139)

To link to this article: <https://doi.org/10.1080/13670050.2024.2433139>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 02 Dec 2024.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

Minority language exposure predicts multilingual receptive abilities in Catalan-Spanish bilingual children

Adriana Soto-Corominas 

Department of English and German Studies, Universitat Autònoma de Barcelona, Bellaterra, Catalunya (Spain)

ABSTRACT

Bilingual profiles may be used to capture the absolute (i.e. not relative) levels of bilinguals' abilities across their two languages. This study investigates the profiles of 146 early Catalan-Spanish bilingual children at the beginning of primary schooling with three goals: (1) to investigate what subgroups of bilinguals exist based on their receptive lexical and grammatical abilities in the two languages, (2) to find the background and linguistic characteristics that distinguish the profiles, and (3) to determine whether or how their profile may relate to their abilities in L3-English. Results showed two main profiles in the sample: a high-abilities group, whose receptive vocabulary and grammar abilities were above the sample mean in both languages, and a low-abilities group, whose results were at or below the mean. The former profile was characterized by more frequent exposure to the minority language Catalan than the latter. In addition, the two profiles exhibited differential performance in their receptive abilities in L3-English, with the high-performing participants scoring significantly higher. Taken together, results demonstrate substantial variation in multilingual skills at the early stages of schooling, and that engaging with the minority language may result in overall multilingual gains due to the benefits conferred by balanced bilingual exposure.

ARTICLE HISTORY

Received 14 November 2023

Accepted 18 November 2024

KEYWORDS

Bilingual profiles; early multilingualism; individual variation; receptive abilities

Introduction

Around 40–50 million Europeans are estimated to speak a regional or minority language (European Commission 2019), most often alongside a majority language. Children brought up in bilingual communities, just like bilingual children raised in monolingual communities, can display ample heterogeneity in their bilingual development (Ezeizabarrena 2012; Gathercole and Thomas 2005; Paradis 2023; Soto-Corominas 2021).

Despite the growing number of studies investigating the sources of individual variation in child bilingual development, studies that directly assess and report on the two languages are relatively scarce, especially in cases of societal bilingualism (i.e. in communities where two or more languages are in contact, where a large proportion of individuals is fluent in these languages). What is more, most of the studies that have investigated the two languages of bilingual children have done so separately, that is, investigating associations between sources of individual differences (e.g. exposure, richness) and development in each language independently (e.g. Gathercole and Thomas 2009; Soto-Corominas 2021; Thordardottir 2014). Studies that have sought to compare abilities across

CONTACT Adriana Soto-Corominas  adriana.soto@uab.cat

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/13670050.2024.2433139>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. 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.

languages have typically done so only in terms of dominance, classifying bilinguals according to whether they show a similar command across languages (*balanced*) or whether their proficiency is asymmetric (*dominant*). While informative, this relative approach does not capture differences in absolute levels of skills (Hoff, Tulloch, and Core 2021). That is, a balanced bilingual could have similarly *low* levels of skills in their two languages compared to other bilinguals.

Bilingual language profiles allow bilingual children to be classified according to their skills across languages. As such, profiles can capture and explain the characteristics that differentiate one subgroup of bilingual children from another. Determining what profiles exist within a group and what predicts such groupings has two main goals. First, knowing what sources of individual variation predict language profiles is of theoretical and applied interest as it informs our knowledge of how language acquisition may be shaped by children's experience. Second, since early language skills predict later outcomes (Hoff 2014; Kieffer 2012), language profiles facilitate the identification of those students who are most at risk for falling behind early on in order to either monitor or provide direct instructional support through targeted intervention (Collins et al. 2014; Su et al. 2024).

Individual differences predicting bilingual profiles

Few studies have investigated individual differences in bilingual development by analyzing bilingual profiles. These studies fall under two categories based on the approach used to create the profiles: either dividing participants into predetermined groups or allowing the data to determine the number and shape of the groups.

The studies by Collins et al. (2014) and Dixon, Wu, and Daraghmeh (2012) applied the former approach. Collins and colleagues tested English-Spanish bilinguals in the Boston (US) area, whereas Dixon, Wu, and Daraghmeh (2012) tested bilingual children in Singapore who spoke English alongside an ethnic minority language (Chinese, Malay, or Tamil). That is, both studies tested participants' English (the majority language) and their heritage minority languages, and divided them into four predetermined groups: (1) high abilities in the two languages, (2) low abilities in the two languages, (3) high abilities in language A and low in language B, and (4) low in language A and high in B. In Collins et al. (2014), the groups were established on the basis of monolingual-normed standardized scores, whereas in Dixon, Wu, and Daraghmeh (2012), it was the sample mean in each language that determined the cutoffs to create the groupings. A common finding across the studies was that language use and socioeconomic status (SES) were predictive of participants' profile. As for language use, Collins et al. (2014) found that children displaying an asymmetric profile (high-low or low-high abilities) had more exposure in the home to the language in which they had higher abilities. On the other hand, in Dixon and colleagues' study, children who spoke both English and the minority language in the home were most at risk for displaying low abilities in the minority language or in the two languages. While it is difficult to draw firm conclusions on the effect of home language on bilingual profiles given the different linguistic contexts in these two studies, their results suggest that there is generally an association between exposure to a language and skills in that language. Regarding SES, Collins et al. (2014) found that children with less educated mothers were more likely to show low abilities in the two languages. Similarly, Dixon, Wu, and Daraghmeh (2012) found that low-SES children were more likely to display low abilities in English or in both languages. That is, both studies suggest that low SES may be predictive of lower bilingual abilities altogether.

Hoff, Tulloch, and Core (2021) avoided pre-defining the potential groupings and instead undertook a data-driven approach to creating the profiles (see also Halpin, Prishker, and Melzi 2021 and Su et al. 2024). In this US-based study, participants took an expressive vocabulary test in English (the majority language) and Spanish (the minority language), and cluster analysis was used to determine what profiles existed in the data given participants' results in the two tests. A total of four profiles emerged, two of which showed an asymmetry between the two languages, English being stronger in both cases. The other two profiles were relatively balanced: one had a moderate total score whereas the other one had high scores in the two languages. Investigating in which dimensions

the four groups differed, authors found that participants in the two balanced profiles lived in homes that were significantly more Spanish-dominant than the other two profiles, and had mothers with more Spanish education.

Hoff et al.'s (2021) results would not be expected if there was a linear relationship between language exposure at home in a language and skills in that language. Given a linear relationship, bilingual children in the US who lived in Spanish-dominant homes would be expected to show a profile with low English abilities and high Spanish abilities, instead of the observed balanced profiles. The authors hypothesized that the use of the minority language in the home fostered *additive* bilingualism. Since they did not find evidence of a trade-off between Spanish and English scores, and English seemed to be acquired relatively uniformly by all children, the use of Spanish at home appeared to boost Spanish abilities, at no cost for English abilities.

Other studies have similarly concluded that strong use of the minority language in the home or with friends is not detrimental to the development of the majority language (Halpin, Prishker, and Melzi 2021; Sagasta 2003). A potential explanation as to why that is the case could be rooted in the notion of *linguistic interdependence*. As part of his hypothesis of linguistic interdependence, Cummins (1979) proposed that a bilingual child's second language (L2) skills are partially predicted by the child's L1 abilities. However, if interdependence applied across the board, positive correlations across languages would be expected, whereas that is far from always the case. In explaining why stronger minority language use is predictive of strong L2 abilities in some children whereas in others there is a null or even negative association, Cha and Goldenberg (2015) invoked Cummins' (1979) *threshold hypothesis*. Specifically, they argued that the transfer of skills from the child's L1 into their L2 would be contingent on the child's degree of bilingualism. That is, the child would need to meet a minimum level of linguistic competence in their L1 for this transfer to occur. Therefore, increased minority language input in the home may benefit L2 skills as long as a minimum threshold level is reached in the L1.

Trilingual development

If linguistic interdependence underlies the profiles of bilingual children with high abilities in their two languages, it follows that these children may have an advantage in L3 acquisition, all else (e.g. access to L3 input, length of L3 exposure) being equal. None of the studies discussed above tested this possibility.

Other studies, however, provide limited evidence that this could be the case. Edele, Kempert, and Schotte (2018) investigated two cohorts of Turkish and Russian heritage language students living in Germany acquiring L3-English at school. Once controlling for potential confounds, the authors found that participants with high abilities in both the heritage language and German performed better in L3-English than (A) those who exhibited low abilities in both languages, and (B) those with low abilities in German and high in the heritage language. However, Edele and colleagues found that the profile that had high German abilities and low heritage language abilities performed on par with the profile with high abilities across languages. Because this study only included samples of bilingual children who spoke a heritage language, their results may have been driven by the fact that only German was the language of schooling. That is, abilities in the language of schooling were the strongest predictor of L3 abilities, while heritage language abilities seemed to play a minor role. Indeed, a study by Sagasta (2003) with bilingual highschoolers whose two languages were regional (Basque and Spanish) and used at school found advantages in L3-English learning for participants with stronger minority language abilities. The bilingual context where the present study took place, Catalonia, is comparable to this latter scenario, as discussed next.

The Catalan context

Catalonia is a bilingual province in Spain where Catalan and Spanish share official status. Importantly, Catalan is the main language of obligatory schooling (ages 6–16) in public and charter (semi-public)

schools (Generalitat de Catalunya 2009). Students are expected to attain fluent oral and written abilities in both Catalan and Spanish by the end of obligatory schooling.

Approximately 81% of Catalonia's population can speak Catalan, while 99.5% can speak Spanish (IDESCAT 2023). This widespread bilingualism is largely due to policy efforts made to protect and promote Catalan, most notably by making it the language of instruction, as noted above. Despite these efforts, only 31.5% of the population report Catalan as their L1, and 36.1% report it as their language of most common use (IDESCAT 2023). Importantly, while Catalan is the language of *schooling* by law, it is not always the language of the *schools*. Observational studies have noted that Spanish is much more widespread than Catalan for communication among peers within the school context, and that Spanish is used to some extent even in teacher-student communication (Plataforma per la Llengua 2019). Given all the above, we follow Extra and Gorter (2009), and many others, in considering Catalan a regional minority language, despite widespread and official bilingualism.

In Catalonia, English is the most prevalent foreign language. However, English abilities are overall low in this region, as only 37.5% of the population over the age of 15 report the ability to speak it to some extent (IDESCAT 2023). English is not widely available in the community outside of tourist hot-spots, and media content (TV shows, movies, videogames) is typically produced in or dubbed into Spanish and, more rarely, Catalan. In school, students are expected to have an intermediate command of English by age 16 (Generalitat de Catalunya 2018).

Catalan and Spanish are closely related Romance languages and display a large degree of lexical and grammatical overlap. On the other hand, English is a Germanic language which, despite a high number of cognates with Catalan and Spanish, is more distant to the two.

Present study

This exploratory study is the first to investigate the bilingual profiles present in a sample of Catalan-Spanish bilingual children. Given that all extant research on bilingual profiles has sampled participants who spoke a heritage language alongside the majority language in largely monolingual communities, the aim of this study is to better understand how certain characteristics of the child and their language use and exposure can influence not just one language specifically, but their bilingual profile in a situation of regional bilingualism. We asked the following research questions:

- (1) What bilingual profiles are present in a sample of Catalan-Spanish bilingual children and how are they described in terms of participants' abilities across languages?
- (2) How do the attested profiles differ with respect to a variety of linguistic and demographic variables (e.g. SES, language use in the home and with friends, richness of the bilingual environments, caregiver language abilities)?
- (3) Do bilingual profiles predict performance in L3-English?

Methods

Participant characteristics

A total of 190 participants (93 females), coming from 14 public or charter schools within the province of Barcelona, participated in this study. Of the initial sample, we excluded the data from participants who did not complete the four target Catalan and Spanish tests (see *Materials* below; $N = 17$) and from participants whose caregivers reported speaking a language other than Catalan/Spanish in the home ($N = 22$) or who had a diagnosis of language delay or autism spectrum disorder ($N = 5$). The final sample comprised 146 participants (71 females). At the time of testing, all participants were starting Grade 1, and had an average age of 6;4 ($SD = 0;4$).

Materials

Background questionnaire

Caregivers who agreed to participate were asked to complete a background questionnaire that collected information on the demographic and linguistic background of participants and their families. At the time of creating the measures for this study, no existing questionnaire captured the specific aspects of interest for a context such as the Catalan one (i.e. a context where all children acquire two languages more or less naturalistically, where they learn a third, foreign, language at school, and where many heritage language children may be acquiring multiple other languages in the home). Therefore, a custom questionnaire was created by adapting questions from other available tools, such as the ALEQ-4 (Paradis et al. 2020) and the LaDoQ (Perpiñán and Soto-Corominas 2021), as well as by formulating new, targeted questions. Caregivers were given the option of completing the questionnaire online, over the phone, or in person. The full questionnaire is available on the OSF repository.

To determine parental levels of education, a proxy for SES in this study, respondents were given nine categories they could choose from, between no formal schooling and a doctorate degree. This variable was subsequently transformed into years of education. Caregivers were also asked to report the participant's age of onset (AO) of acquisition for each of the languages, as well as the context where exposure to each language began (either at home, at school, or in the community).

Since all participants in this sample only used Catalan and/or Spanish in the home, we asked about their language use in the family and with their friends using a relative 1–5 scale: 1 = Catalan always/Spanish never (100%/0%); 2 = Catalan most of the time/Spanish hardly ever (75%/25%); 3 = Catalan and Spanish to a similar extent (50%/50%); 4 = Catalan hardly ever/Spanish most of the time (25%/75%); 5 = Catalan never/Spanish always (0%/100%). The variable *Language use with family* averaged the values for the language used between the participant and their mother, father, older siblings, younger siblings, uncles, cousins, and maternal and paternal grandparents. The variable *Language use with friends* averaged the values for friends from school and friends outside of school.

Caregivers were asked to report the number of hours per week participants spent engaging in three language-rich activities in each of the three languages: reading (including homework, parent–child joint reading), listening (to TV shows, videogames, cartoons), and partaking in extracurricular activities. From caregivers' responses to the questionnaire, an input diversity score was calculated. The questionnaire included 14 human and non-human sources that could potentially provide input in each of the languages (namely: mothers, fathers, older siblings, younger siblings, cousins, uncles, maternal grandparents, paternal grandparents, friends in school, friends outside of school, babysitter, TV/videogames, reading material, and extracurriculars). We then calculated a proportion score (0–1) for each participant and language, where 1 would indicate that all given sources provided input in that language and 0 would indicate that none did. Importantly, a participant could plausibly get a diversity score of 1 in all languages as long as all input providers used all languages with the participant. When sources were not applicable (e.g. the participant had no siblings), those relevant input sources were removed from the denominator.

Teacher survey

While Catalan is by policy the language of instruction, children may make different linguistic choices to communicate with peers while at school (Plataforma per la Llengua 2019). Teachers were asked to indicate, in an interview, the linguistic typology of the class. They were given a variety of options to describe the nature of the classroom, out of which only three were selected: *the majority of my students speak Catalan in the home and in the classroom*, *the majority of my students speak Spanish in the home and in the classroom*, and *there is a similar proportion of students who speak Catalan and Spanish in the home and in the classroom*. This variable was used to consider the input of the wider community that participants may be exposed to outside of the home.

Vocabulary tests

In order to test receptive vocabulary abilities, we employed three tests with the same design. Children were presented with an array of four pictures and were asked to point to the picture that best matched the word spoken by the experimenter. For English we used the *Peabody Picture Vocabulary Task – 5th edition* (Dunn 2019). For Spanish we used the *Test de vocabulario en imágenes PPVT-III* (Dunn, Dunn, and Arribas 2010), which had been normed on Peninsular Spanish.

Since Catalan does not have a standardized vocabulary test, we adapted the *Test de Vocabulario en Imágenes Peabody* (TVIP; Dunn et al. 1986), which was originally developed for Latin American Spanish. The TVIP was initially compared to the Spanish PPVT-III and revealed 30 items in common. To alleviate effects of task reinforcement, an effort was made to replace these items in the Catalan version of the task, while the PPVT-III was left intact for its administration in Spanish.

The replacement items were of the same part of speech as the target item, of a similar log frequency according to the SUBTLEX Catalan corpus and, whenever possible, were close in meaning. For example, *cow* was replaced by *pig*, and not by *pocket*, despite sharing a similar frequency with both words. Since words of a similar frequency in the corpus may pose differential challenges for children, three primary school teachers were queried to ascertain whether the original item and the proposed alternative word presented a similar degree of difficulty. Once items were agreed upon, a black-and-white line drawing in the style of the original test was found or commissioned to replace the original item.

The three vocabulary tests were discontinued when a given ceiling criteria was reached. For the Catalan and English tests, which did not contain item sets, the ceiling was reached after 6 mistakes in a group of 8 items. For Spanish, the ceiling was reached after 8 mistakes in a set of 12 items. The Cronbach's alpha of internal consistency was .92 for the Catalan test, .93 for the Spanish test, and .98 for the English test.

Grammar tests

The three tests used to evaluate receptive grammatical abilities again had identical protocols. Participants heard a sentence and were asked to choose which picture out of four best matched the sentence. For Spanish we used the *Comprensión de Estructuras Gramaticales* (Mendoza et al. 2005). For English we used the *Test for Reception of Grammar 2* (Bishop 2003). Using 80 items, these tests assess the comprehension of 20 different grammatical structures of varied complexity, with each structure being tested by four items. To shorten the task, we reduced the items to 40 by randomly selecting two items for each structure. Catalan has no test of receptive grammar. We constructed the Catalan grammar test using the 40 items that had been eliminated from the Spanish test, translating them into Catalan, without the need for further adaptation. As such, the Catalan and Spanish tests had comparable difficulty. The three tests were discontinued after 10 consecutive mistakes or null responses.

Since these grammar tests inevitably rely on vocabulary skills, participants were given a targeted vocabulary test to determine whether they knew the nouns, verbs, and adjectives included in the 40 grammar items. Those words that participants did not know prior to the test were presented and tested twice before the grammar test. The Cronbach's alpha was .71 for the Catalan test, .64 for the Spanish test, and .88 for the English test.

Procedures

Ethical considerations

The protocols for this study were revised and approved by the ethics board at the home university (protocol number MUL-2021-01). Teachers at participating schools shared the invitation to participate in the study with caregivers. Caregivers who agreed to participate were sent the background questionnaire. Only those children whose questionnaires were submitted participated in the data collection sessions.

Data collection

Data was collected at school during the school day, during times chosen by teachers. Participants were removed from class and tested individually in a quiet space. Participants completed three data collection sessions, one for each language. Only one language was tested in a given session. The three sessions happened around one week apart. Testing sessions lasted for a maximum of 50 minutes including breaks. The language for the first and second session was counterbalanced between Catalan and Spanish across schools. English was always tested in the third session.

Data analysis

To identify linguistic profiles within our sample (research question 1) we employed k-means clustering, a method that partitions datasets by allowing for observations that are closer to one another to be grouped together (Pham, Dimov, and Nguyen 2005). Given the differences in scale between the tests, raw scores were transformed to z-scores in order to find the optimal number of centers. We determined the optimal number of centers by considering the within-cluster sum of squares and the average silhouette method (Saputra, Saputra, and Oswari 2020). Both methods found the optimal number of centers in the data to be two.

Research question 2 asked about the differences between the clusters. Two-tailed independent-samples t-tests were employed to compare the clusters with regard to continuous variables, and cumulative link models with the package *ordinal* (Christensen 2022) were used to compare the two clusters with respect to ordinal variables (e.g. parental language skills). Given the small sample size in this study, trends in contrasts that did not reach statistical significance are discussed in the text as we believe that they may inform directions for future research.

To address question 3, in order to determine whether there was an effect of profile after controlling for SES and L3 experiential variables that could affect participant performance in L3-English, we ran two mixed-effects negative binomial models with the package *lme4* (Bates et al. 2015). The dependent variables of these models were the total score from the English vocabulary test, in the first model, and the score from the grammar test in the second. The predictors of these models were *Profile* together with SES and other L3 experiential variables, as outlined in the Results section. Numerical variables were scaled and centered, and factor variables were dummy-coded. There was one random intercept for *School* to control for variability arising from participants coming from 14 schools. Model diagnosis was performed using the DHARMa package (Hartig 2022).

Results

Sample characteristics

Participant characteristics appear in Table 1. The mean number of years of education for mothers and fathers in the sample was close to 16 years, which corresponds to a bachelor's degree. This mean is expected given the population, as 53% of those living in Catalonia between the ages of 30–34 have post-secondary studies (IDESCAT 2023).

As shown in Table 1, the mean AO of the three languages was early. In fact, most participants could be considered simultaneous bilinguals in Catalan and Spanish. For English, the mean AO was just below age 3. The early AO for English should be understood together with the context of first exposure to the language (see Table 2). While the home/family was the context where most of the participants were first exposed to Catalan or Spanish, most participants were first exposed to English at school or in extracurricular activities. Given that 94.52% of our sample had been schooled by age 3, it is unsurprising that they were exposed to English at such an early age at school. The percentage of participants schooled by age 3 in the current sample is very close to the population average, given that census data indicates that 94.30% of 3-year-olds in Catalonia were schooled in the academic year 2020–2021 (IDESCAT 2023).

Table 1. Participant characteristics.

Variable	Mean	SD	Range
Maternal education (in years)	15.45	2.61	10–22
Paternal education (in years)	14.75	2.86	10–22
Catalan AO (in months)	7.04	12.67	0–60
Spanish AO (in months)	4.05	9.76	0–36
English AO (in months)	29.58	15.97	0–78
Language use in the family (1–5)	2.97	1.49	1–5
Language use with friends (1–5)	3.22	1.21	1–5
Weekly hours of Catalan reading	2.95	3.22	0–20
Weekly hours of Spanish reading	2.74	3.77	0–20
Weekly hours of English reading	0.75	1.42	0–8
Weekly hours of Catalan TV/videogames	1.60	2.21	0–15
Weekly hours of Spanish TV/videogames	4.80	4.29	0–20
Weekly hours of English TV/videogames	1.31	2.61	0–20
Weekly hours of Catalan extracurricular activities	1.91	1.44	0–6
Weekly hours of Spanish extracurricular activities	0.75	1.26	0–6
Weekly hours of English extracurricular activities	0.67	0.76	0–3
Input diversity in Catalan	0.72	0.28	0–1
Input diversity in Spanish	0.69	0.28	0.08–1
Input diversity in English	0.12	0.09	0–0.37

As seen in [Table 1](#), the mean value for *Language use in the family* corresponds to the mid-point of the scale with notable spread in the data. Given the large spread, the mean should not be interpreted as indicating that most participants lived in families where the two languages were spoken. Instead, it indicates that the sample included participants who were more frequently exposed to Catalan and participants who were more exposed to Spanish in the home. A similar statement can be made for the *Language use with friends*.

In terms of the frequency of language-rich activities and input diversity in each language, there was substantial variation in the sample. However, and as would be expected, the frequency of English-rich activities and the input diversity in this language were generally lower than for the two societal languages.

Descriptive test results

The descriptive statistics for the language tests appear in [Table 3](#) below. Since the tests had different ranges, these are indicated in the first column. Vocabulary tests are not directly comparable to one another since the possible range of scores differed across the three languages. The grammar tests, however, were more comparable across languages since they each had 40 items and included similar syntactic structures.

Bilingual profiles

In clustering participant performance to identify linguistic profiles in the sample, we considered the four tests corresponding to Catalan and Spanish. As explained in the Data analysis section, the optimal number of clusters (i.e. profiles) was two. We then proceeded to plot the two profiles with respect to their performance on the vocabulary ([Figure 1](#)) and grammar tests ([Figure 2](#)). It should be noted that [Figures 1–2](#) present raw scores rather than the z-scores used for clustering

Table 2. The distribution of participants (in percentages) according to their context of first exposure to each of the three languages.

Context	Catalan	Spanish	English
Home/family	71.23%	80.14%	10.27%
Community	0.69%	8.22%	0%
School/extracurriculars	28.08%	11.64%	89.73%

Table 3. Descriptive test results, as raw scores.

Test	Mean	SD	Range
Catalan vocabulary (out of 125)	57.68	12.15	30–90
Spanish vocabulary (out of 192)	73.29	14.92	35–117
English vocabulary (out of 240)	21.77	16.64	3–94
Catalan grammar (out of 40)	27.90	4.24	12–38
Spanish grammar (out of 40)	28.88	3.83	15–38
English grammar (out of 40)	8.77	7.34	1–35

in order to enhance the interpretability of the data. As shown in [Figures 1 and 2](#), one of the two clusters performed generally above the mean in both languages for vocabulary and grammar, while the other group performed at or below the mean. This is corroborated by the descriptive and inferential statistics regarding test performance in [Table 4](#). For simplicity, we refer to the two clusters as *high abilities* and *low abilities*, respectively.

Differences between bilingual profiles

To answer research question 2, which asks about the background and linguistic differences between the profiles, we present these characteristics separately for the high- and low-abilities profiles in [Table 5](#), together with the *p*-value of a t-test comparing the two profiles. Whenever the test is significant, a Cohen’s *d* is also provided in the table. As shown in [Table 5](#), mothers of participants in the high-abilities profile were significantly more educated than those of the low-abilities profile. There was a trend for fathers of participants in the high-abilities profile to also be more educated, though this comparison did not reach statistical significance. The two profiles differed significantly on several linguistic dimensions: *Language use in the family*, *Language use with friends*, *Weekly hours of Catalan reading*, and *Input diversity in Catalan*. For all these variables, the high-abilities profile exhibited higher proportions or frequencies for Catalan than the low-abilities profile. Finally, there was a trend for participants in the high-abilities profile to have more *Weekly hours of Catalan extra-curricular activities*.

In addition to the continuous variables in [Table 5](#), we investigated whether the profiles differed on any categorical dimensions that could potentially affect language development, namely in parental language skills and classroom linguistic typology.

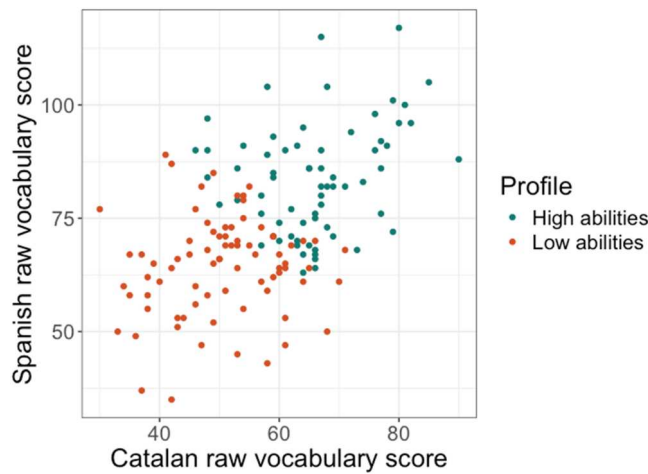


Figure 1. Participants’ raw scores in the vocabulary tests in Catalan and Spanish. Individual participants are colored according to their profile.

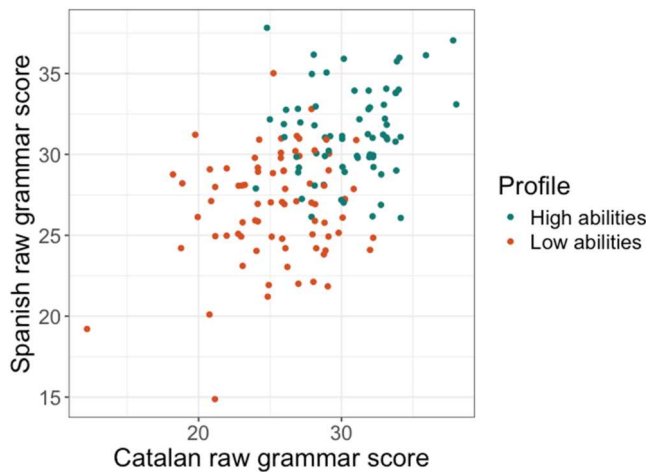


Figure 2. Participants’ raw scores in the grammar tests in Catalan and Spanish. Individual participants are colored according to their profile. Points appear jittered (i.e. slightly shifted randomly) to avoid overlap.

Respondents were asked to rate their language skills in the three languages using an ordinal scale with descriptors: *Non-existent, Limited, Basic, Intermediate, Advanced, and Highly advanced* skills. [Figure 3](#) shows the distribution of parental skills for the three languages, further subdivided by profile.

Since self-rated language abilities are ordered categories, we ran three ordinal models, one for each language, where the outcome variable was *Caregiver self-rated skills*, and the only predictor was *Profile* (high vs. low abilities). The full output of these models appears in Appendix 1. The models found no differences between clusters for Catalan or Spanish skills. With regards to caregiver English skills, the model found that participants in the high-abilities cluster were more likely to have caregivers with higher self-rated English skills ($p = .030$).

Finally, we consider the linguistic typology of the classroom, visualized in [Figure 4](#). An ordinal model found a tendency for participants classified in the low-abilities profile to be in majority-Spanish classrooms ($p = .078$). This trend, however, was not significant. The full output of this model appears in Appendix 2.

L3-English skills by profile

The third research question asks whether the two profiles would perform differently in L3-English. Indeed, two t-tests found that there were significant differences for the two English tests. For vocabulary, the high-abilities profile ($M = 25.75$; $SD = 17.16$) outperformed the low-abilities profile ($M = 18.13$; $SD = 15.39$) significantly: $p = .007$, $d = .469$. For grammar, the high-abilities profile ($M = 10.77$; $SD = 8.10$) similarly outperformed the low-abilities profile ($M = 6.77$; $SD = 5.91$): $p = .002$, $d = .564$.

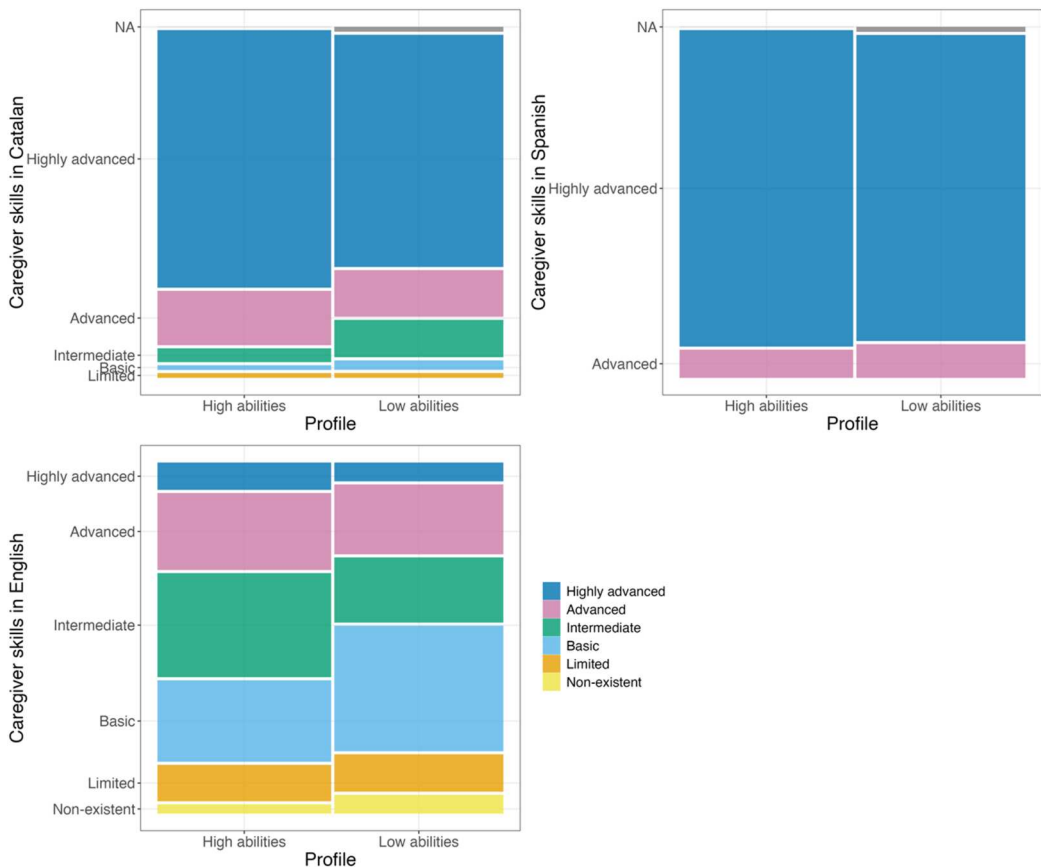
To explore whether the effect of profile remains once SES and English exposure are controlled for, we fit two mixed-effects negative binomial models: one for vocabulary and one for grammar. The

Table 4. Descriptive statistics for the Catalan and Spanish tests according to participant profile, and significance value resulting from an independent-samples t-test, together with Cohen’s d effect size.

	High abilities			Low abilities			t-test, Cohen’s d
	Mean	SD	Range	Mean	SD	Range	
Catalan vocabulary	65.90	9.40	46–90	50.50	9.39	30–71	$p < .001$; $d = 1.644$
Spanish vocabulary	83.50	12.20	63–117	64.40	10.9	35–89	$p < .001$; $d = 1.655$
Catalan grammar	30.70	3.06	24–38	25.50	3.61	12–32	$p < .001$; $d = 1.538$
Spanish grammar	31.30	2.81	26–38	26.80	3.32	15–35	$p < .001$; $d = 1.471$

Table 5. Profile characteristics and *p*-value resulting from a t-test. Cohen's *d* included when t-test was significant.

Variable	High abilities			Low abilities			t-test, Cohen's <i>d</i>
	Mean	SD	Range	Mean	SD	Range	
Maternal education	16.11	2.49	10–22	14.87	2.59	10–22	$p = .004$; $d = 0.490$
Paternal education	15.27	2.79	10–22	14.32	2.87	10–22	$p = .051$
Catalan AO (in months)	5.84	11.56	0–36	8.11	13.57	0–60	$p = .278$
Spanish AO (in months)	4.51	9.45	0–36	3.65	10.06	0–36	$p = .595$
English AO (in months)	30.04	16.18	0–78	29.18	15.89	0–77	$p = .747$
Language use in the family (1–5)	2.61	1.43	1–5	3.29	1.48	1–5	$p = .005$; $d = 0.471$
Language use with friends (1–5)	2.91	1.13	1–5	3.49	1.22	1–5	$p = .004$; $d = 0.489$
Weekly hours of Catalan reading	3.62	3.79	0–20	2.37	2.50	0–10	$p = .023$; $d = 0.393$
Weekly hours of Spanish reading	2.62	3.62	0–20	2.84	3.91	0–20	$p = .722$
Weekly hours of English reading	0.65	1.11	0–6	0.84	1.64	0–8	$p = .399$
Weekly hours of Catalan TV/videogames	1.46	1.87	0–7	1.73	2.48	0–15	$p = .446$
Weekly hours of Spanish TV/videogames	4.96	4.54	0–20	4.67	4.09	0–20	$p = .681$
Weekly hours of English TV/videogames	1.07	1.65	0–8	1.52	3.22	0–20	$p = .279$
Weekly hours of Catalan extracurricular activities	2.13	1.29	0–6	1.71	1.54	0–6	$p = .074$
Weekly hours of Spanish extracurricular activities	0.66	1.14	0–4	0.83	1.35	0–6	$p = .425$
Weekly hours of English extracurricular activities	0.76	0.77	0–3	0.60	0.74	0–3	$p = .184$
Input diversity in Catalan	0.79	0.25	0–1	0.65	0.29	0–1	$p = .002$; $d = 0.519$
Input diversity in Spanish	0.65	0.28	0.09–1	0.73	0.27	0.08–1	$p = .101$
Input diversity in English	0.13	0.09	0–0.37	0.11	0.08	0–0.30	$p = .383$

**Figure 3.** Parental skills in Catalan (top left), Spanish (top right), and English (bottom left).

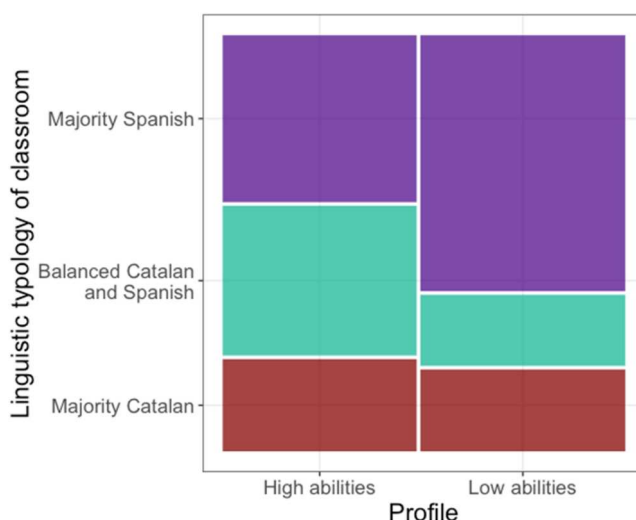


Figure 4. Linguistic typology of the classroom according to profile.

dependent variables were the test scores, and they were predicted by *Profile* together with other variables that could affect English abilities: *English AO*, *Weekly hours of English extracurricular activities*, *Weekly hours of English TV/videogames*, *Weekly hours of English reading*, *Input diversity in English*, and *Parental self-rated English skills*. Finally, *Maternal years of education* was added as an additional co-variate in order to control for the effect of SES.

Once the effect of all the co-variables was controlled for, in the vocabulary model, the effect of profile was no longer significant. In the grammar model, the effect of profile reached statistical significance ($p = .015$). The two models predicted lower scores for participants in the low-abilities profile. The full output for the two models appears in Appendix 3.

Discussion

This exploratory study was the first to employ a bilingual language profile approach in the pursuit of expanding knowledge on the sources of variation of bilingual abilities in cases of regional bilingualism. Conceptualizing a child's bilingual abilities as the single outcome of bilingual development has the advantage of capturing the absolute levels of skills in the two languages simultaneously, which allows us to describe groups of bilinguals and the demographic/linguistic variables that characterize them. In addition to examining the bilingual profiles present in our Catalan-Spanish bilingual sample, we probed into the relation between the profiles and performance in L3-English.

Number and nature of bilingual profiles

The studies reviewed in the introduction investigated the abilities of bilingual children who spoke a heritage language in the home and were acquiring English as a majority language (Collins et al. 2014; Dixon, Wu, and Daraghmeah 2012; Hoff, Tulloch, and Core 2021). As opposed to Collins et al. (2014) and Dixon, Wu, and Daraghmeah (2012) but in line with Hoff, Tulloch, and Core (2021), we did not predetermine the number of groups. However, unlike Hoff, Tulloch, and Core (2021), the cluster analysis found that the optimal number of profiles to capture our data was two: a profile that, on average, had high abilities in the two languages and a profile that had low abilities. That is, unlike Hoff, Tulloch, and Core (2021), no asymmetric profiles were found (i.e. high in Catalan and low in Spanish, or the reverse). Given that our sample differs in two crucial aspects from that of Hoff and

colleagues', this discrepancy appears well-founded. First, the participants in this study were born and raised in a bilingual community where both languages share official status and are present in the community and at school. This already creates a more balanced situation than the ones in prior studies. Second, and perhaps more importantly, Catalan and Spanish are two Romance languages that share a high degree of cross-linguistic similarity – more so than English and Spanish. It is likely that high cross-linguistic similarity fosters the existence of symmetrical bilingual profiles, at least where receptive abilities are concerned. We discuss the potential effects of cross-linguistic similarity further in the next section.

Differences between profiles

As the two profiles divided participants into *high* and *low abilities*, investigating what variables distinguished the two profiles entailed looking into the variables that predicted better or worse bilingual performance overall. To this extent, we investigated a range of demographic and linguistic variables.

The two profiles differed significantly on many dimensions. First, participants in the high-abilities profile had parents with more education. This is far from the first study to find that higher parental education is associated with higher linguistic outcomes; an association that has been noted both for monolingual and bilingual children (Collins et al. 2014; Hart and Risley 2003; Hoff, Tulloch, and Core 2021). However, the difference in means of parental education between the two profiles was relatively small (approximately one year). As such, it seems unlikely that this small difference could have had a meaningful impact on participants' bilingual development.

The other variables on which the two profiles differed were directly related to linguistic experience. Participants in the high-abilities profile used significantly more Catalan with their family and friends, engaged in more hours of reading activities in Catalan, and received Catalan input from more diverse sources. In addition, there were two other trends that, while not statistically significant, were consistent with these previous results: participants in the high-abilities profiles spent more time taking part in extracurricular activities in Catalan and were less likely to be in Spanish-dominant classrooms than participants in the low-abilities profile. While the variables were diverse, the findings were consistent: participants in the high-abilities profile exhibited patterns of language exposure and use that largely involved more Catalan.

Minimally, these results evidence that when it comes to receptive abilities in Catalan and Spanish, languages do not take away from each other. Higher abilities in one language are positively associated to higher abilities in the other language. In addition, engaging more frequently with the minority language appears to confer advantages not only for the development of Catalan but of multilingual abilities more broadly. At least superficially, these results align with Hoff, Tulloch, and Core (2021) and Sagasta (2003), who found that increased use of the minority language boosted bilingual skills overall by virtue of strengthening minority language skills at no cost for the majority language.

The remaining question is what the precise mechanism is that underlies the observed association between increased exposure to Catalan and high multilingual abilities overall. One possibility is the hypothesis of *linguistic interdependence*, which broadly suggests that proficiency in an L2 is partly dependent on someone's L1 proficiency (Cummins 1979). As predicted by this hypothesis, the bilingual participants in our sample with more exposure to the minority language seem to develop stronger receptive skills in Catalan, which then transfer into the majority language in the shape of stronger receptive Spanish skills. This transfer is facilitated by the majority status of Spanish, which is present in the community, in the media, and at school. Therefore, children with higher Catalan skills have enough exposure to Spanish to allow them to deploy their Catalan receptive skills in the development of their Spanish skills. On the other hand, children's whose exposure in the home is mostly limited to Spanish may not be exposed to enough Catalan input to allow them to draw on their L1 (Spanish) receptive skills in the development of their L2 (Catalan). While this explanation

would fit the observed data, it would also predict the existence of a third profile: one with high Spanish skills and low Catalan skills. This profile was not attested in the data.

A second possibility is that the two clusters simply reflect differences in an unobserved construct, such as verbal or non-verbal intelligence. In this regard, the high-abilities cluster would be formed by participants with higher cognitive skills who are thus better language learners overall, with the reverse holding for the low-abilities cluster. This possibility cannot be tested directly, as participants were not administered tests that measured these constructs. However, a systematic pattern of by-cluster differences emerged in this study, by which the high-abilities grouping had consistently more exposure (or richness) in Catalan than the other group. If groupings were solely driven by general cognitive abilities, this systematicity would not be expected.

The explanation that seems to fit the data best is that increased bilingual abilities are the result of relatively balanced bilingual exposure. Given that Spanish is the majority language, richer Catalan environments yield more balanced bilingual exposure overall. The idea that balanced bilingualism may be advantageous is far from new. Studies have reported that bilinguals with balanced exposure may have advantages in executive control performance as a function of engaging more frequently with language selection or inhibition (e.g. Bosma, Blom, and Versloot 2017). As a result, Catalan-Spanish bilingual children with more balanced exposure to their two languages may be more efficient language learners, displaying higher abilities overall. A potential aid for participants is the high cross-linguistic similarity between Catalan and Spanish. The two languages share a large number of cognates and many syntactic properties (e.g. Bel 2001). Given that cross-linguistic similarities are known to be facilitative and to promote positive transfer (Ringbom 2007), we argue that balanced Catalan-Spanish exposure affords bilinguals increased opportunities for this positive transfer to take place, resulting in higher bilingual performance overall.

Profile performance on L3-English

Even though children were tested at the onset of primary school, the high-abilities profile already exhibited significantly higher receptive vocabulary and grammar abilities in English than the low-abilities profile, despite the two profiles being comparable in terms of the variables that could directly impact L3-English linguistic experience. The only exception were self-reported caregiver English skills. In this regard, participants in the high-abilities profile were found to have parents who rated themselves as having higher English abilities. However, this is unlikely to affect participants' L3-English skills directly, given that caregivers in this sample used Catalan and/or Spanish exclusively with participants. Instead, this difference may reflect the somewhat higher SES of the families in high-abilities group, indexed by the higher parental education.

In order to determine whether the profile-based difference remained once other factors that could influence L3 abilities were controlled for, two models were run in order to predict vocabulary scores, on the one hand, and grammar scores, on the other. In the former model, the presence of the SES and L3 experiential covariates swamped the effect of profile. In the case of grammar, the effect of profile remained significant.

In explaining why the high-abilities profile may exhibit an advantage in L3-English, at least in grammar, there are multiple potential explanations that fit the observed results. First, the linguistic interdependence hypothesis coupled with the threshold hypothesis seem to provide a fitting answer (Cummins 1979). Those bilinguals that have attained a higher level of L1-L2 skills are better able to transfer their abilities into L3 learning. The second explanation, in line with Edele et al.'s findings (2018), is that better performance in L3-English is due to the fact that the high-abilities profile has higher abilities in Catalan, the main language of schooling. The development of English may be faster for those learners who have a stronger command of the language learning occurs in. Disentangling these two competing explanations is not possible in Catalonia, where public and charter schools adhere to the use of Catalan as the language of schooling. Lastly, the findings for L3-English are also compatible with the potential by-profile differences in cognitive skills put forth

above (i.e. that participants in the high-abilities profile have higher (verbal) intelligence than participants in the other profile). That is, if differential cognitive abilities are indeed at the root of the profile separation, then the better performance of the higher-abilities profile in L3-English would be expected.

Limitations, conclusions and implications for educators

This study was not without limitations. First, we could not use standardized tests in Catalan given the lack of suitable instruments. Even though the internal validity criteria of the adaptations made for this study were very high, these tests have not gone through the validation process that comes with standardization procedures. A second limitation is that only receptive abilities were tested as part of this study. Future research should determine if results hold when productive skills are examined.

Despite the limitations of this study, there are some noteworthy findings. As opposed to previous studies investigating the nature of bilingual profiles, we only encountered two profiles in our Catalan-Spanish bilingual sample: children with high abilities in the two languages, and children with low abilities. In the classroom, a child that displays low receptive abilities in both Catalan and Spanish may experience difficulties engaging with academic content and meeting learning expectations. Therefore, this finding suggests that bilingual children whose receptive abilities in Catalan or Spanish are weaker in comparison to that of their peers may be good candidates for close monitoring, as low receptive abilities in one of the languages are likely to indicate low receptive bilingual abilities overall, rather than language dominance.

While engaging more with the minority language was argued to be advantageous for overall multilingual development in this sample due to the benefits of balanced exposure, it may not be feasible or even advisable to recommend that children use more Catalan with friends or family to boost bilingual abilities. Alternatively, educators should ensure that children have sufficient exposure to Catalan in the school context, and that they have the possibility of engaging in meaningful interaction with others (both instructors and children alike) in Catalan at school and during extracurricular activities.

Acknowledgements

I would like to thank all participants as well as their families and schools for their time and effort. I am also very thankful to all the research assistants who collaborated in this project.

Data availability statement

Supplemental data for this article can be accessed online at <https://osf.io/aym98/>.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101027137.

Notes on contributor

Adriana Soto-Corominas is an Assistant Professor in the Department of English and German Studies at Universitat Autònoma de Barcelona, Catalonia (Spain). Her research is primarily concerned with early multilingual development in children and youth.

ORCID

Adriana Soto-Corominas  <http://orcid.org/0000-0002-0489-9098>

References

- Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67 (1): 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bel, A. 2001. *Teoria lingüística i adquisició del llenguatge: anàlisi comparada dels trets morfològics en català i en castellà*. Barcelona: Institut d'Estudis Catalans.
- Bishop, D. V. M. 2003. *The Test for Reception of Grammar, Version 2 (TROG-2)*. Oxford: Pearson.
- Bosma, E., E. Blom, and A. Versloot. 2017. "Language Balance and Cognitive Advantages in Frisian-Dutch Bilingual Children." In *Bilingualism and Minority Languages in Europe: Current Trends and Developments*, edited by M. C. Parafita Couto and F. Lauchlan, 141. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Cha, K., and C. Goldenberg. 2015. "The Complex Relationship Between Bilingual Home Language Input and Kindergarten Children's Spanish and English Oral Proficiencies." *Journal of Educational Psychology* 107 (4): 935. <https://doi.org/10.1037/edu0000030>
- Christensen, R. H. B. 2022. Ordinal - Regression Models for Ordinal Data. R package version 2022.11–16. <https://CRAN.R-project.org/package=ordinal>.
- Collins, B. A., E. E. O'Connor, C. Suárez-Orozco, A. Nieto-Castañón, and C. O. Toppelberg. 2014. "Dual Language Profiles of Latino Children of Immigrants: Stability and Change Over the Early School Years." *Applied Psycholinguistics* 35 (3): 581–620. <https://doi.org/10.1017/S0142716412000513>
- Cummins, J. 1979. "Linguistic Interdependence and the Educational Development of Bilingual Children." *Review of Educational Research* 49 (2): 222–251. <https://doi.org/10.3102/00346543049002222>
- Dixon, L. Q., S. Wu, and A. Daraghme. 2012. "Profiles in Bilingualism: Factors Influencing Kindergartners' Language Proficiency." *Early Childhood Education Journal* 40 (1): 25–34. <https://doi.org/10.1007/s10643-011-0491-8>
- Dunn, L. M. 2019. *The Peabody Picture Vocabulary Test*. 5th ed. Bloomington: NCS Pearson.
- Dunn, L. M., L. M. Dunn, and D. Arribas. 2010. *Test de vocabulario en imágenes PPVT-III*. 3rd ed. Madrid: TEA Ediciones.
- Dunn, L. M., E. R. Padilla, D. E. Lugo, and L. M. Dunn. 1986. *Manual del examinador para el Test de Vocabulario en Imágenes Peabody*. Circle Pines: American Guidance Service.
- Edele, A., S. Kempert, and K. Schotte. 2018. "Does Competent Bilingualism Entail Advantages for the Third Language Learning of Immigrant Students?" *Learning and Instruction* 58:232–244. <https://doi.org/10.1016/j.learninstruc.2018.07.002>
- European Commission. 2019. *The Teaching of Regional or Minority Languages in Schools in Europe*. Luxembourg: Publications Office of the European Union.
- Extra, G., and D. Gorter. 2009. "Regional and Immigrant Minority Languages in Europe." In *Handbook of Language and Communication: Diversity and Change*, edited by M. Hellinger and A. Pauwels, 15–52. Berlin: De Gruyter.
- Ezeizabarrena, M. J. 2012. "The (in) Consistent Ergative Marking in Early Basque: L1 vs. Child L2." *Lingua. International Review of General Linguistics. Revue Internationale De Linguistique Generale* 122 (3): 303–317. <https://doi.org/10.1016/j.lingua.2011.11.009>
- Gathercole, V. C. M., and E. M. Thomas. 2005. "Minority Language Survival: Input Factors Influencing the Acquisition of Welsh." In *Proceedings of the 4th International Symposium on Bilingualism*, edited by J. Cohen, K. T. McAlister, K. Rolstad, and J. MacSwan, 852–874. Somerville, MA: Cascadia Press.
- Gathercole, V. C. M., and E. M. Thomas. 2009. "Bilingual First-Language Development: Dominant Language Takeover, Threatened Minority Language Take-up." *Bilingualism: Language and Cognition* 12 (2): 213–237. <https://doi.org/10.1017/S1366728909004015>
- Generalitat de Catalunya. 2009. *Llei 12/2009, del 10 de juliol, d'educació*. Diari Oficial de la Generalitat de Catalunya. https://educacio.gencat.cat/web/contenut/home/departament/normativa/normativa-educacio/lec_12_2009.pdf.
- Generalitat de Catalunya. 2018. *El model lingüístic del sistema educatiu de Catalunya*. Departament d'Ensenyament. <https://educacio.gencat.cat/ca/departament/publicacions/monografies/model-linguistic>.
- Halpin, E., N. Prishker, and G. Melzi. 2021. "The Bilingual Language Diversity of Latino Preschoolers: A Latent Profile Analysis." *Language, Speech, and Hearing Services in Schools* 52 (3): 877–888. https://doi.org/10.1044/2021_LSHSS-21-00015
- Hart, B., and T. R. Risley. 2003. "The Early Catastrophe: The 30 Million Word Gap by Age 3." *American Educator* 27:4–9.
- Hartig, F. 2022. DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.6. <https://CRAN.R-project.org/package=DHARMA>.
- Hoff, E. 2014. *Language Development*. Belmont: Cengage Learning.
- Hoff, E., M. K. Tulloch, and C. Core. 2021. "Profiles of Minority-Majority Language Proficiency in 5-Year-Olds." *Child Development* 92 (5): 1801–1816. <https://doi.org/10.1111/cdev.13591>

- IDESCAT. 2023. Basic indicators of Cagalonía. Accessed August 2023. <https://www.idescat.cat/indicadors/>.
- Kieffer, M. J. 2012. "Early Oral Language and Later Reading Development in Spanish-Speaking English Language Learners: Evidence from a Nine-Year Longitudinal Study." *Journal of Applied Developmental Psychology* 33 (3): 146–157. <https://doi.org/10.1016/j.appdev.2012.02.003>
- Mendoza, E., G. Carballo, J. Muñoz, and M. D. Fresneda. 2005. *Test de Comprensi3n de Estructuras Gramaticales*. Madrid: TEA Ediciones.
- Paradis, J. 2023. "Sources of Individual Differences in the Dual Language Development of Heritage Bilinguals." *Journal of Child Language* 50 (4): 793–817. <https://doi.org/10.1017/S0305000922000708>
- Paradis, J., A. Soto-Corominas, X. Chen, and A. Gottardo. 2020. "How Language Environment, age, and Cognitive Capacity Support the Bilingual Development of Syrian Refugee Children Recently Arrived in Canada." *Applied Psycholinguistics* 41(6) (6): 1255–1281. <https://doi.org/10.1017/S014271642000017X>
- Perpiñán, S., and A. Soto-Corominas. 2021. "Indirect Structural Crosslinguistic Influence in Early Catalan–Spanish Bilinguals in Adulthood: Predicate Selection in Catalan Existential Constructions." *Applied Psycholinguistics* 42 (6): 1463–1502. <https://doi.org/10.1017/S0142716421000308>
- Pham, D. T., S. S. Dimov, and C. D. Nguyen. 2005. "Selection of K in K-means Clustering." *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 219 (1): 103–119. <https://doi.org/10.1243/095440605X8298>
- Plataforma per la Llengua. 2019. *Estudi sociolingüístic als patis d'escoles i instituts de zones urbanes de Catalunya*. https://www.plataforma-llengua.cat/media/upload/pdf/estudi-socilinguistic-patis-web-ok_1630936308.pdf.
- Ringbom, H. 2007. *Cross-linguistic Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters.
- Sagasta, M. P. 2003. "Acquiring Writing Skills in a Third Language: The Positive Effects of Bilingualism." *International Journal of Bilingualism* 7 (1): 27–42. <https://doi.org/10.1177/13670069030070010301>
- Saputra, D. M., D. Saputra, and L. D. Oswari. 2020. "Effect of Distance Metrics in Determining k-Value in k-Means Clustering Using Elbow and Silhouette Method." In *Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)*, 341–346. Atlantis Press.
- Soto-Corominas, A. 2021. "Acquisition of Quantified Partitivity in Catalan-Spanish Bilingualism: Influence from Child-Level and Language-Level Factors." *Linguistic Approaches to Bilingualism* 11 (6): 873–908. <https://doi.org/10.1075/lab.19016.sot>
- Su, P. L., R. Rojas, J. de Villiers, R. Golinkoff, and A. Iglesias. 2024. "Receptive Dual-Language Profiles in Spanish–English Bilingual Preschoolers." *Journal of Educational Psychology* 116 (1): 102–122.
- Thordardottir, E. 2014. "The Typical Development of Simultaneous Bilinguals." *Input and Experience in Bilingual Development* 13:141.