ANUARI DE FILOLOGIA. ESTUDIS DE LINGÜÍSTICA (Anu.Filol.Est.Lingüíst.)
12/2022, pp. 109-124. ISSN: 2014-1408. DOI: 10.1344/AFEL2022.12.5

SPANISH LEXICAL ACCESSIBILITY AND VOCABULARY ACQUISITION: CATALAN DOMINANT BILINGUALS AND SPANISH MONOLINGUALS COMPARED

MARIO RUIZ MORENO
Johannes Gutenberg-Universität Mainz
mruizmor@uni-mainz.de
ORCID: 0000-0002-9504-1045

ABSTRACT

In this study it was investigated whether Catalan dominant bilinguals from Girona performed differently on a Spanish vocabulary test in comparison to Spanish monolinguals from Toledo. The test measured both lexical knowledge and lexical accessibility. The test was administered to two groups of bilinguals and monolinguals: one aged 5-6 and another aged 8-9. The monolinguals outperformed the bilinguals at both ages, but most of the Catalan dominant bilinguals fell within monolingual norm at age 8-9, despite the fact that at group level the differences were still observable.

KEYWORDS: bilingualism, Spanish, Catalan, lexical accessibility, vocabulary, acquisition.

ACCESSIBILITAT LÈXICA I ADQUISICIÓ DE VOCABULARI EN CASTELLÀ: COMPARACIÓ ENTRE BILINGÜES DOMINANTS EN CATALÀ I MONOLINGÜES EN CASTELLÀ

RESUM

Accessibilitat lèxica i adquisició de vocabulari en castellà: comparació entre bilingües dominants en català i monolingües en castellà. En aquest estudi s'analitza si els bilingües dominants en català residents a Girona i els monolingües en castellà residents a Toledo es comporten de manera diferent en un test de vocabulari castellà. El test emprat mesura tant el coneixement lèxic com l'accessibilitat lèxica, i es va administrar a dos grups de monolingües i bilingües en cada territori; un format per parlants de entre 5 i 6 anys, i un altre constituït per parlants de 8 i 9 anys. Els monolingües van superar als bilingües en tots dos grups d'edat, però la majoria dels bilingües dominants en català d'entre 8 i 9 anys van entrar dins la norma monolingüe, malgrat que les diferències entre tots dos grups encara eren observables.

PARAULES CLAU: bilingüisme, castellà, espanyol, català, accessibilitat lèxica, adquisició de vocabulari.

1. INTRODUCTION

The aim of this paper is to compare the performance of four groups of speakers in a Spanish vocabulary test: one group of Spanish monolinguals aged 5-6, another of Spanish monolinguals aged 8-9, and two more of age-matched Catalan dominant bilinguals. Because of the nature of the test (described in 3.2),

Data de recepció: 19/III/2021
Data d'acceptació: 3/V/2021
Data de publicació: desembre 2022

the results give us information both about the size of these speakers' lexicon and their accessibility to it.

It is somewhat controversial to compare the performance of bilingual speakers with that of monolinguals because, in the past, the results of this type of studies have been interpreted as a linguistic deficit and used to support monolingual linguistic ideologies; the aim of this paper could not be any further from that. The goal pursued here is to understand the linguistic differences between monolinguals and bilinguals, which is not only a legitimate one from a scientific point of view, but also one with important social and educational implications. Knowing what is realistic, typical, and possible in language proficiency among bilingual speakers should not only be born in mind when designing educational policies for bilingual communities but is also crucial to help parents decide in which language they want to speak to their children. In this regard, Rhys & Thomas explicitly state that

> Welsh-medium teaching or bilingual education that focuses mainly (or solely) on increasing exposure to, and acquisition of, Welsh is only viable if English proficiency is met independently. (Rhys & Thomas 2012: 650)

In other words, Welsh-medium teaching will encounter resistance if English proficiency is not met. The opinion that minority language proficiency should not be attained at the expense of that of the majority language is not necessarily shared by all parents in bilingual territories, but it seems safe to assume that is the prevalent one; therefore, for those who seek to promote bilingual education, comparing the linguistic performance between monolinguals and bilinguals is not only legitimate, but also necessary so that no parents are afraid that their children will have linguistic deficits and can choose what language(s) they want to speak at home.

In the Spanish context, it is relatively recurrent to find articles on the press (mostly in the conservative one) reporting a supposedly lack of Spanish linguistic competence among Catalan children, but such remarks are frequently based on anecdotic evidence and generally from a monolingual perspective. These articles generally report some sentences with examples of Catalan transfer, most frequently in the domain of lexicon, or examples of momentaneous problems of lexical accessibility. Politicians in Catalonia usually counterargued that the grades in *Lengua castellana y literatura* ('Spanish language and literature') for the university entrance exams are actually higher in Catalonia than in many other monolingual *comunidades autónomas* (≈ federal states). However, the fact that these grades in *Lengua castellana y literatura* are higher than in many other monolingual *comunidades autónomas* cannot be claimed as evidence of a higher linguistic performance. To begin with, because both the syllabi and the University entrance exams are designed and implemented by each regional government, which invalidates all comparisons between the different *comunidades autónomas*. Secondly, because neither oral

lexical accessibility nor lexical knowledge are measured in these exams. Hopefully, the data here reported will contribute to hold this necessary discussion in more scientific terms.

In order to achieve this goal, a test to measure lexical accessibility was designed and implemented in two schools, one in the city of Girona, located in a dominant Catalan area, and one in Numancia de la Sagra, located in the province of Toledo, which in turn is located in the *comunidad autónoma* of Castilla-La Mancha, i.e., a Spanish monolingual area. The research analyses the data collected for an unpublished master thesis (Ruiz Moreno 2010), but both works present major modifications: firstly, some participants included in the original study have been now excluded in order to guarantee intragroup homogeneity, because in their families languages other than Catalan and Spanish were also spoken; secondly, there have been major changes in the way the data have been analysed and discussed (because of the different number of participants considered and because this literature review is updated and much more critical).

2. LITERATURE REVIEW

The studies that will be reviewed in this section were carried out with bilinguals living in different contexts with social bilingualism in Europe, namely the Irish-English, Welsh-English, Frisian-Dutch, Basque-Spanish, and Catalan-Spanish contexts. Several of these studies measured the lexical competence in the two languages of the bilinguals, but in this study only the Spanish lexicon of the bilinguals was analysed; therefore, the review of each of these language combinations will concentrate on the data of the stronger language. By stronger language it is here meant the language with the higher number of speakers is meant, not the stronger language in terms of proficiency at individual level.

Parsons & Lyddy (2009) carried out a cross-sectional vocabulary study with 254 students aged 6, 8, and 10, as well as a longitudinal one with 84 students in Count Galway (Ireland). The students were divided in four different groups depending on the school type they attended to. The Gaeltacht school group, i.e., the group with higher Irish competence and input, scored significantly below the other groups at the age of 8 in the English test, but not at the ages of 6 and 10. The results of their longitudinal study are in line with those of Stephens, who also carried out a study with different types of bilinguals, including a Gaeltacht group with 19 students and concluded that

> monolingual children, children in immersion education, and children from Gaeltacht areas of Ireland showed equivalent skills in their levels of English vocabulary. (Stephens 2013: 293)

Lastly, O'Toole & Hickey (2016) compared the total vocabulary, i.e., including both languages, of Irish dominant bilingual toddlers (8 to 36 months) with that of American ones and concluded that their total vocabulary was similar to that of American children of the same age and that they knew more words in Irish than in English.

In Wales, at least two important studies have analysed both the lexical knowledge in Welsh and English. Gathercole & Thomas (2009) conducted a study of receptive vocabulary study in English with 240 children in Wales whose ages ranged from 3 to 11 and concluded that the acquisition of English was not hindered even in the case of bilinguals from only-Welsh-speaking homes who also attended only-Welsh schools and lived in areas like Gwynedd, where 69 % of the population speaks Welsh. Specifically, they noted that the effect of home language had been neutralised by age 7 for the English tests, but also that at age 9 school-language was a significant factor. Additionally, they also administered a vocabulary test to adults and the scores of those who spoke only English at home were higher than those who spoke only Welsh at home, showing a "near-significantly better performance". Rhys & Thomas (2012) also conducted a study on receptive vocabulary in English in Wales, for which they recruited a sample of 207 speakers from North Wales between the ages 7 and 11 divided into linguistic profiles (socioeconomic status was controlled for). The English monolinguals and the L1 English-L2 Welsh bilinguals performed significantly better on the English tests than the simultaneous bilinguals and the L1 Welsh bilinguals even at age 11.

Like Spanish and Catalan, but unlike Welsh or Irish and English, Frisian and Dutch are closely related languages. In the context of Frisian and Dutch bilingualism, several investigations have also been carried out. Dijkstra *et al*. (2016) recruited 91 participants from families where predominantly only one language was spoken (58 and 33 with Frisian and Dutch as their home language, respectively); their age range was within 2,6 and 4 years. The L1-Dutch speakers scored higher than the L1-Frisian speakers on the Dutch receptive vocabulary test. The lag was small, but constant over time. Nonetheless, the factor home language was not statistically significant. On the contrary, what the authors labelled as "outside home exposure" reached significance. As for the Dutch productive vocabulary, the L1 Dutch speakers outperformed the L1 Frisian participants, but this time home language did reach significance, whereas outside home exposure did not (the *p*-value was .06, though). These authors also refer to a work published by Van Ruijven (2006), in which Frisian-Dutch bilinguals performed similarly in Dutch compared to other children in the rest of the Netherlands.

In the Basque Country, Etxebarria recruited large samples of participants for two studies (1996 & 2014) that measured the Spanish lexical production among Basque-Spanish bilinguals and Spanish monolinguals. Her methodology is very similar to the one employed in this study (section 3), since the

participants were required to name as many items as possible for a given thematic within a time-limit of two minutes. Her participants, however, are considerably older than in the rest of the studies (secondary and even pre-university education students) and were divided into two groups: those for whom Basque was the medium of instruction and those who received their lessons in Spanish. For the latter group it is pretty safe to assume that we are mostly or exclusively dealing with Spanish monolinguals, but in the former group there are students with very different home language backgrounds (some of them may speak exclusively Basque at home, whereas others exclusively Spanish and, yet have opted for the only-Basque model). The results of the first study (Etxebarria 1996), with a sample of 245 students, show that the Basque-Spanish bilinguals scored slightly lower than those who followed an only-Spanish program in the Basque Country. The author also compared these scores with those of other monolingual speakers tested in Madrid or Las Palmas for other studies but with the same methodology; in both regions the monolinguals scored between the Basque-Spanish bilinguals and the Spanish monolinguals residing in the Basque Country, which seemingly proves that factors other than amount of input are to be considered, namely, socio-cultural and economic status. The same author also obtained almost identical results in a posterior study (Etxebarria 2014) with 345 participants.

Lastly, Blas Arroyo & Casanova Ávalos (2003) also applied the same methodology to 246 scholars from the province of Castelló in the Valencian Community, where Catalan and Spanish are spoken (their exact age is not reported). When the study was conducted, and as noted by the authors, the amount of Catalan input at school of these students may have varied considerably, but typically they would mostly receive instruction in Spanish. These students were divided into different groups depending on the input they received at home. Home language turned out to be statistically significant, although had a moderate effect. It must be noted, however, that, as the authors point out, the educational level of the Spanish-speaking parents was higher than that of the Catalan-speaking ones. In turn, those who heard both Catalan and Spanish at home scored better than those who were spoken to exclusively in Spanish, which apparently cannot be accounted for by the educational level of the parents, since it was also lower than in the families who spoke exclusively Spanish.

Globally, the studies here reviewed show that there is an initial linguistic gap for the majority language between bilinguals and monolinguals, but also that, given sufficient time, all bilinguals can catch up with the monolinguals to a large extent or even completely in all the European contexts here reviewed. This was true even for those bilinguals living in areas where the majority language is not dominant and even receiving most of their instruction in their minority language. The age at which the gap between the monolinguals and bilinguals is reduced may vary considerably and most likely depends on a combination of

different factors: hours of instruction in the minority language, non-school linguistic practices and input, and the degree of language-relatedness with the stronger language, among others.

To conclude, it is also worth remembering that most of the studies here reviewed did not control for the effect of the socio economic and/or the educational level of the parents, which has definitely to be taken into account for the interpretation of the data. In some contexts, the bilinguals typically have a higher educational and economic status, which help reduce the effect of their reduced input because of their bilingualism; in others, the bilinguals typically have a lower educational and economic status, which may enlarge the differences in the majority language caused in relation to the monolinguals.

In this regard, the studies of Etxebarria (1996 & 2014) are particularly illustrative; her Spanish monolinguals residing in the Basque Country performed better on the Spanish test than her Basque-Spanish bilinguals, which suggests an effect of the factor bilingualism, but her Basque-Spanish bilinguals performed better than the monolinguals tested in other regions with the same methodology. Another study in which the amount of input cannot explain alone the linguistic outcome of the speakers is that of Mägiste (1986), who conducted a study at a German school in Sweden in which the L1 German students who received extensive instruction in German scored better in different Swedish tests than other age-matched students from state schools in Sweden (i.e., with Swedish as the sole medium of instruction). Lastly, Pérez Pereira (2004) also carried a study in which language input cannot alone account for the results. In his study, the Galician lexical knowledge of 706 speakers of different ages (from 8 to 30 months) was tested. Astonishingly, those who received bilingual input at home scored slightly better in the Galician tests than those who were addressed to exclusively in Galician.

3. EMPIRICAL STUDY

In this section, the methodological aspects (3.1 Participants and 3.2 Test) will be first described; thereafter, the results (3.3), which include both quantitative analyses (3.3.1) and some qualitative remarks (3.3.2), will be presented.

3.1. Participants

A total of 65 participants were recruited and distributed into four groups:

Group	Participants	Age	Schooled in
Very young bilinguals	15	5-6	Girona
Very young monolinguals	21	5-6	Numancia de la Sagra
Young bilinguals	14	8-9	Girona
Young monolinguals	15	8-9	Numancia de la Sagra

TABLE 1. Participants

As can be seen in Tables 1, 2, and 3, the age-matched groups are reasonably well balanced in terms of number of participants but differ in the educational level of their progenitors (four parents in Numancia de la Sagra refused to provide this information).

		Number of parents by educational level	
		Non-university education %	University education %
Very young bilinguals	Mother	33	67
	Father	47	53
Very young monolinguals	Mother	95	5
	Father	95	5
Young bilinguals	Mother	0	100
	Father	71	29
Young monolinguals	Mother	92	8
	Father	100	0

TABLE 2. Percentage of mothers and fathers with and without university education

		Number of parents by educational level	
		Non-university education	University education
Very young bilinguals	Mother	5	10
	Father	7	8
Very young monolinguals	Mother	19	1
	Father	19	1
Young bilinguals	Mother	0	14
	Father	10	4
Young monolinguals	Mother	11	1
	Father	12	0

TABLE 3. Number of mothers and fathers with and without university education

Whereas families in which at least the mother had university studies clearly predominate in Girona, only one of the families from Numancia de la Sagra reported a progenitor with university education. Besides, no socioeconomic data were collected, but given the differences in the educational level of the parents and the general socioeconomic differences between these two cities, it is more than safe to presume that they must be of great magnitude.

As for the linguistic profile of these speakers, it must be said that both parents were native in Spanish in the case of the monolingual speakers, and all Catalan families included in the study reported to speak Catalan exclusively or at least 80 % of their time at home. Moreover, Numancia de la Sagra is located in Castilla-La Mancha, a monolingual region, whereas Girona is part of Les Comarques Gironines, a territory where Catalan is the most common choice among progenitors (figure 1) and is also the dominant language. In its *Informe de política linguística 2018*, the Catalan Government (Generalitat de Catalunya

2018: 21) reports that Catalan was the ordinary language for 54.1 % of the respondents from Les Comarques Gironines, whereas Spanish was the main language for only of the 29 % (the rest either reported other languages or a balanced use of both Spanish and Catalan). These data are also more favourable to Catalan than they were in 2013; that year, 51.5 % reported to use mainly Catalan, whereas 30.8 % mainly Spanish.

Summing up, the results that will be reported in 3.3 correspond to a very specific type of bilinguals: those who clearly qualify as Catalan dominant bilinguals. These bilinguals make (almost) exclusive use of Catalan at home, attend classes in which Catalan is the medium of instruction, and live in a Catalan dominant area; thus, they are by no means representative of the whole bilingual population of Catalonia. In this regard, it is worth remembering that Catalonia has a population of around 7.67 million people, of which circa 3.23 inhabit Barcelona and its metropolitan area, where Spanish is predominant (figure 2). Indeed, according to the data reported by the Catalan Government, in this region Spanish dominated over Catalan in 2018; 48.6 % of the respondents reported to use primarily Spanish, whereas 36.1 % primarily Catalan.

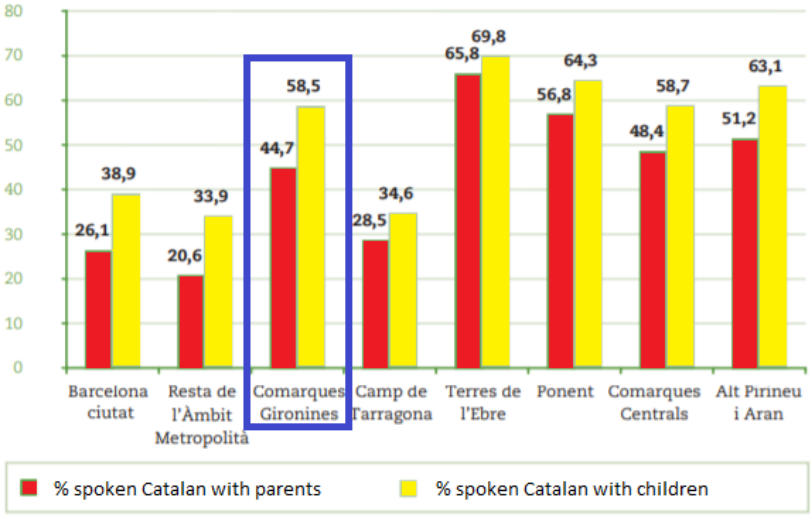

FIGURE 1. Use of Catalan at home in 2008, modified from Generalitat de Catalunya (2011: 90)

3.2. Test

The test was administered orally and individually to the participants, who had previously been told that they would receive the visit of somebody to play a fun game with them and that he would ask them some questions. The game was called "The fantastic journey to India" and consisted of two human-shaped pieces and a board decorated with colourful and gaudy ornaments and with eye-catching pictures (emblematic Indian sites, animals, and other aspects of Indian culture).

Since this study pursued to measure lexical accessibility in Spanish, the initial idea was to explain the instructions in Spanish to all children in order to

help them activate their Spanish knowledge; however, this procedure had to be changed after testing a couple of very young bilinguals because they clearly had problems at understanding the rules of the game in Spanish. The instructions were repeated in Catalan for those first participants; the rest of the bilinguals were directly addressed in Catalan, except for those aged 8-9, who had no problem whatsoever at understanding the instructions in Spanish.

The participants were told that they were going to play a board game whose goal was to reach the final square, for which they had a minute to provide as many answers as possible for the questions they would receive. The game consisted of eight different categories (presented in this order): colours, animals, things you can eat, garments, things you can find in a classroom, things you can see in the sky, professions, and states of mind / feelings. For each category, the student was asked to name as many items as possible. Those children who did not know what they were supposed to answer for a given category were told one example in Catalan; this extra help was only occasionally necessary for the categories of "states of mind", and "professions".

The bilinguals were told that their wrong answers would not be penalised and were explicitly encouraged to provide as many answers as possible, since otherwise they would not reach the final square, where they could claim their reward (different packs of stickers). Their reward was shown in advance to serve as visual stimulus to the child and, before beginning the actual test, a trial round with a different lexical category was played with all children so that all of them were familiarised with the game. Most of the children understood the rules of the game perfectly and seemed to have a very good time during the test. To avoid feelings of failure, the researcher made sure that, regardless of their actual performance, every child advanced the necessary number of square to complete the game and receive their reward.

Some of the answers were problematic because they showed signs of Spanish linguistic knowledge which was not possible to capture with the binary methodology (right or wrong) here used to count the number of valid answers. For example, a speaker uttered **conejil*, a mixed form created from *conejo* and *conill* ('rabit' in Spanish and Catalan respectively). These forms were not accepted as valid answers. In addition, some of the answers consisted in definitions rather than words; for example, in the category "professions" some children tended to say things like *el que vende pescado*, 'the one who sells fish', or *el que trabaja en la pescadería*, 'the one who works in a fishmonger's', instead of *pescadero*, 'fishmonger'. In those cases, the children were asked how you call any person who works doing that / in that place; most of the times this explanation sufficed, but responses such as *Antonio* or *Andreu* were not rare. In this and other similar cases, only the word referring to the profession was accepted as a valid answer (i.e., *pescadero* 'fishmonger' in this example). Lastly, a third source of problematic answers was the use of non-prototypical answers; for instance, "Pope" for the category of professions, or "the angels", "my grandma and my

dog, who are dead", "butterflies", "UFOs", "aliens" or "Martians" as response to "things that you can see in the sky". Virtually all these non-prototypical answers were accepted for both groups. This kind of problem only occurred when testing the very young speakers (both monolinguals and bilinguals), because the young participants, aged 8-9, had a much more adult-like view of the world.

3.3. Results

Although this study was not designed to perform qualitative analyses, some remarks in this direction will be given in 3.3.2 after the core part of this section, 3.3.1, in which the quantitative differences between the bilingual and monolingual speakers will be reported.

3.3.1. Quantitative analyses

As can be seen in figures 2 and 3, the participants followed the expected trend and the monolinguals were able to name more items than the bilinguals, both at ages 5-6 and 8-9.

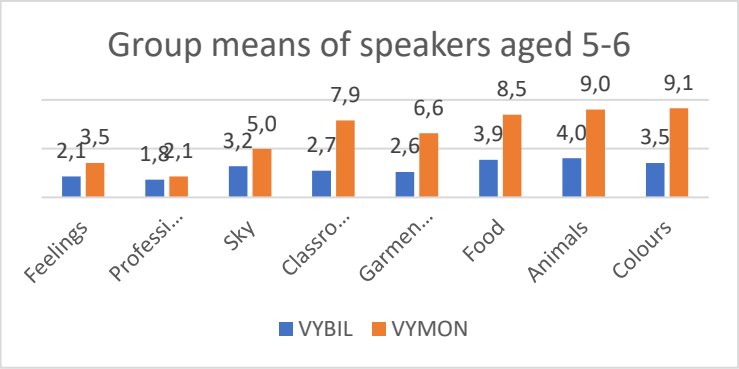

FIGURE 2. Group means of speakers aged 5-6

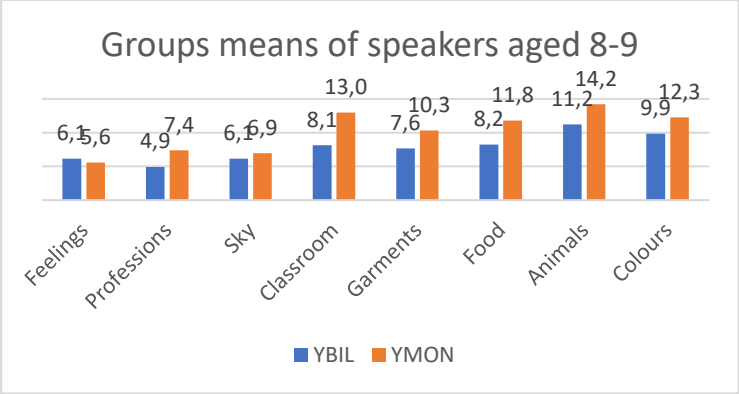

FIGURE 3. Group means of speakers aged 8-9

As can be inferred from comparing figures 2 and 3, there is a clear reduction of the performance gap between monolinguals and bilinguals as the speakers grow older. A battery of two-tailed Mann-Whitney tests was carried

out for each lexical category between the bilinguals and their age-paired monolinguals. Differences between the two groups at ages 5-6 were highly significant for each category (p <.001), whereas at ages 8-9, there were important differences in the p-values: colours (.021), animals (.041), food (.004), garments (.068), classroom (<.001), sky (.227), professions (.016), and feelings (.478). Following the recommendations made for the field of linguistics in Larson-Hall (2010), α was set at 0.1, which yields the result that differences between monolinguals and bilinguals were also significant in all categories at ages 8-9 except for the categories "things you can see in the sky" and "feelings"; for this latter category, the bilinguals actually attained a higher mean.

Given the heterogeneity typically found among bilinguals, it is worth going beyond group means and pay attention to individual means. Table 4 shows each participant's mean score across all lexical categories, whereas the lowest and highest scores per group are given in Table 5, which also contains the difference between the ratio between the maximum and the minimum score. As can be seen, the highest score among 5-6 bilinguals is 7.8 times higher than that the lowest one in the same group, but at ages 8-9 the same difference has been drastically reduced to a ratio of just 1.8. On the contrary, there is just a negligible reduction of the ratio between the highest and lowest means for the group of monolinguals between the same ages.

5-6-BIL	5.4	4.4	4.4	4	3.9	3	2.2	2.1	2	1.7	1.6	1.4	0.9	0.9	0.7					
5-6-Mon	7.6	7.3	6.6	6.2	6	5.7	5.6	5.6	5.6	5.6	5.6	5.2	5.1	5.1	4.9	4.5	4.3	4	3.5	3.4	3.2
8-9-BIL	8.5	8.3	8.2	7.9	7.8	7.8	7	6.4	6.4	6.1	5.7	5.4	5.1	4.7							
8-9-Mon	12.6	11.3	10.8	10.1	9.2	9.2	9	9	8.8	8.3	8.3	7.5	7	7	5.6						

TABLE 4. Individual means per group

Group		Min	Max	Max/MinRatio
5-6-BIL		0.7	5.4	7.8
8-9-BIL		4.7	8.5	1.8
5-6-Mon		3.2	7.6	2.4
8-9-Mon		5.6	12.6	2.3

TABLE 5. Lowest and highest scores per group

The intragroup variability in performance of the 8-9 aged bilinguals is drastically reduced in comparison to that of the 5-6 bilinguals. This research was not designed to account for intragroup differences; however, it is possible to speculate that the interaction between language aptitude and input seems to be a good candidate to account for these changes. As reviewed in Ruiz Moreno (2020: 80-90), language aptitude is probably much more important than previously assumed, especially when the learning conditions are difficult. In this line of thought, one can speculate that speakers with a high language aptitude will shine more when the amount of input is reduced, and that, conversely, differences in language aptitude will become partially masked when there is a considerable amount of input. The fact that the older bilinguals

have received more input than the younger bilinguals seems to support this interpretation.

Beyond group means, it is also interesting to see what happens at the individual level. The aim of this study was to determine whether these bilinguals have comparable lexical skills in Spanish to those of the monolinguals; therefore, the mean score of the monolinguals was not taken as reference and all bilinguals who performed equal or better than the lowest monolingual score were considered to perform within monolingual norm. The rationale behind this decision is that all the monolingual speakers included in this study have, by definition, monolingual lexical skills, regardless of their performance, since they have been brought up monolingually in a monolingual territory, and none had been diagnosed with a learning deficit. In other words, taking the mean score of the monolingual group as the reference would be such a severe reference that even many monolinguals would fail to perform within monolingual ranges.

Interestingly, following these criteria, only a third of the 5-6 aged bilinguals fell within monolingual norm; however, 78.6 % of the bilinguals at ages 8-9 performed within monolingual ranges. In other words, even though that at age 8-9 the bilinguals still perform significantly lower at group level than the monolinguals in almost all lexical categories, most of these dominant Catalan bilinguals have mean scores that fall within monolingual norm.

A couple of remarks on these results must be made, however. Firstly, as noted in 3.1, there were important differences between the monolinguals and the bilinguals both in the educational level of the progenitors and in their socioeconomic status. Had not the monolinguals a lower socioeconomic status and educational parental background, the differences between the monolinguals and the bilinguals would have presumably been higher. Secondly, this was not a longitudinal study, and the socioeconomic differences between the monolinguals and the bilinguals were higher among the speakers aged 8-9 than among those aged 5-6, which means that in a longitudinal study the reduction rate for the performance gap would have probably taken place at a slower pace.

3.3.2. Qualitative remarks

It is out of the scope of this study to provide a thorough quantitative analysis of the participants' answers; however, it is worth briefly commenting some of the invalid answers because they provide important clues for a better understanding of how bilingual children produce and access vocabulary in their non-dominant language; especially when, like in the Catalan – Spanish combination, the two languages are closely related.

Given this linguistic similarity, it was relatively common that many participants hesitated whether, for example, the Spanish word for 'computer'

was *ordenador* or **ordinador*. Even though *ordinador* is actually the Catalan for 'computer', it was possible to tell that the child thought it possible to be also a Spanish word because the pronunciation had been adapted to Spanish, i.e.[oɾdinaˈðo̞ɾ] instead of [uɾdinəˈðo̞], which would be the normal pronunciation in their Catalan variety. In those cases, in which a word was uttered with a Spanish pronunciation, but the real Spanish word had not been uttered, the answer was counted as invalid.

Interestingly, they not only converted Catalan words into pseudo-Spanish ones by adapting them to the Spanish phonetic rules, but also derived Spanish words from Catalan words applying the correspondences typically found in many cognates. Thus, following the patterns found in pairs of words such as Catalan *baix* [baʃ] and Spanish *bajo* [ˈbaχo] or Catalan *caixa* [ˈka.ʃə] and ˈSpanish *caja* [ˈka.xa], a child in the 5-6 aged group derived the word **llangardejo* from Catalan *llangardaix* instead of producing Spanish *lagarto* 'lizard'. Similarly, a kid aged 8-9 derived the pseudo-Spanish word **calajo* from Catalan *calaix* 'drawer'; in both cases, they are aware that there Catalan and Spanish share many cognates and when they do not know a word or cannot access to it, they rely on their Catalan knowledge and apply their unconscious linguistic knowledge to derive Spanish words from the Catalan ones. Logically, this linguistic device can only be noticed when the children fail to produce the target word. Further examples are provided in Table 6 (none of these pseudo-Spanish words was accepted as valid).

Child's answer in pseudo-Spanish	Catalan word	Spanish word	English word
*Vienos	Venus	Venus	Venus
*fustero	fuster	carpintero	carpenter
*escalfadores	escalfadors	calentadores	leg warmers
*dolfín	dofí	delfín	dolphin
*goanes	guants	guantes	gloves
*zol	sol	sol	sun
*zabates	sabates	zapatos	shoes
*tristo	trist	triste	sad
*colomes	coloms	palomas	doves
*núbolos	nubols	nubes	clouds
*coja	cuixa	muslo	thigh
*michones	mitjons	calcetines	socks
*astruz	astruç	avestruz	ostrich
*cambrero	cambrer	camarero	waiter
*bocabalado	bocabadat	boquiabierto	open-mouthed
*exploziones	explosions	explosiones	explotions
*grogo	groc	amarillo	yellow
*oselos	ocells	pájaros	birds

TABLE 6. Sample of pseudo-Spanish words uttered during the research

Finally, it is also worth noting that many children, especially the younger ones, answered sometimes in Catalan and then in Spanish; for example, they

would first whisper the Catalan word *cadira* and then loudly utter the Spanish word *silla* 'chair', which can be interpreted as a sign that these bilinguals first activate the concept through the lexicon of their dominant language and then access their lexical repertoire in their weaker language.

4. DISCUSSION

The results reported in 3.1 are in line with the existing literature on bilingual acquisition in similar European contexts in that the initial Spanish linguistic skills of these Catalan dominant bilinguals clearly lagged behind those of the monolinguals, but only three years later these differences have been drastically reduced and most bilinguals at age 8-9 fell within monolingual norm (although at group level there were still differences). It is nonetheless necessary to note anew that the monolinguals in this study had a lower socioeconomic status and educational parental background. In all likelihood, if the monolinguals and bilinguals had the same status and parental background, they bilinguals would take longer to fall within monolingual norm. Besides, it is also important to highlight that these bilinguals are not representative of the whole bilingual population in Catalonia, but rather of the type of bilinguals who receive a relatively small amount of input in Spanish, since they speak (almost) exclusively Catalan at home and the language of instruction at school is Catalan, apart from living in a Catalan dominant area.

In the study of Gathercole and Thomas (2009) with Welsh-English bilinguals, the effect of home language had been neutralised for the English tests at age 7, but at age 9 those who received some instruction in English still outperformed those who did not in the English vocabulary test, and even the adult speakers presented some minor differences. Rhys & Thomas (2012) also found differences on English performance even at age 11 among different schoolers in Wales. The very young Frisian-Dutch bilinguals of Dijkstra *et al*. (2016) had a lower performance in Dutch than the Dutch monolinguals, but not the Frisian-Dutch bilinguals in Van Ruijven (2006). Lastly, Etxebarria (1996 & 2014) carried out two studies with Basque-Spanish bilinguals who were considerably older than those in the other studies here reported; the mean scores of these bilinguals were still lower than that of the Spanish monolinguals, although did not reach significance. All in all, it is observable that the exact time at which the bilinguals catch up with, or come very close to, the monolinguals inevitably varies from one context to another, but in all likelihood also depends on the complexity of the linguistic task and on individual circumstances like language aptitude. However, it can also be observed that, at least for some linguistic tasks, it is not realistic to expect a fully comparable performance at group level between the monolinguals and the bilinguals. Since monolingual speakers can be viewed as highly specialised speakers of a language at the expense of learning no other, it would be rather

surprising that we would not find even the slightest difference between bilingual and monolingual speakers. These differences may be visible even in adulthood, when the bilinguals have had time enough to receive vast amounts of input in their weaker language, but that does not necessarily mean, of course, that these bilinguals have a linguistic deficit in their weaker language; it simply reflects that they are not specialised on a single language as the monolinguals are, and, yet, they are fully functional even in their weaker language.

Studies like this are of interest for the general public and for language-policymakers and should be more often carried out and expanded in the future, since only with a continuous scientific assessment can language policies in bilingual regions be adapted to ensure a high linguistic competence in both languages. Additionally, having reliable data of the students' linguistic abilities will enrich the legitimate political discussion of to what extent Catalan and Spanish need to be used at school.

5. ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to all the participants, their parents, and the staff of the two schools where this study was conducted. They did not only make the study possible, but also a very nice and fun experience. Secondly, I would like to thank Virginia Gathercole and Enlli Thomas (my MA dissertation supervisor and my MA coordinator) for all the valuable advice received for the elaboration of the master dissertation, from which this study has originated. Finally, I would like to thank an anonymous reviewer for a series of critical comments that have helped to improve the paper.

BIBLIOGRAPHY

BLAS ARROYO, J. L. & CASANOVA ÁVALOS, M. (2003), "La influencia de la lengua materna en algunos valores de la disponibilidad léxica en el español de una comunidad bilingüe", in *Adquisición, enseñanza y contraste de lenguas, bilingüismo y traducción*, Doval Reixa, I. & Pérez Rodríguez, M. R. (eds.), Vigo, Servicio de publicacións da Universidade de Vigo, pp. 49-61.

DIJKSTRA, J. *et al.* (2016), "The role of the majority and minority language input in the early development of a bilingual vocabulary", *Bilingualism: Language and Cognition*, 19(1), 191-205.

ETXEBARRIA, M. (1996), "Disponibilidad léxica en escolares del País Vasco: variación sociolingüística y modelos de enseñanza bilingüe", *Revista española de lingüística*, 26, 301-325.

ETXEBARRIA, M. (2014), "Lexical variation and bilingual education in the Basque country", *Spanish in Context*, 11(1), 50-75.

GENERALITAT DE CATALUNYA. *Enquesta d'usos lingüístics de la població 2008: anàlisi. Volum I. Les llengües a Catalunya: coneixements, usos, transmissió i actituds lingüístics*

[online]. Solé, J. (ed.), 2011 [Access date: 07/10/2022]. Available at: <https://llengua.gencat.cat/permalink/2e24ce66-5384-11e4-8f3f-000c29cdf219>.

GENERALITAT DE CATALUNYA. *Informe de política lingüística 2018* [online]. [Access date: 07/10/2022]. Available at: <https://llengua.gencat.cat/web/.content/documents/informepl/arxius/IPL-2018.pdf>.

GATHERCOLE, V. & THOMAS, E. M. (2009), "Bilingual first-language development: dominant language takeover, threatened minority language take-up", *Bilingualism: Language and Cognition*, 12, 213-237.

LARSON-HALL, J. (2010), *A guide to doing statistics in Second Language research using SPSS*, New York & London, Routledge.

O'TOOLE, C. & HICKEY, T. M. (2016), "Bilingual acquisition in a minority context: using the Irish-English Communicative Development Inventory to track acquisition of an endangered language", *International Journal of Bilingual Education and Bilingualism*, 20, 2, 146-162.

PARSONS, C. & LYDDY, F. (2009), "Learning to read in Irish and English: A comparison of children in Irish-medium, Gaeltacht and English-medium schools in Ireland", *Final report to An Chomhairle um Oideachas Gaeltachta & Gaelscolaíochta*, 1-32.

PÉREZ PEREIRA, M. (2004), "Efectos do bilingüismo na adquisición da linguaxe dos nenos pequenos: um estúdio con nenos galegos", in *Proceedings from the Second International Symposium on Bilingualism,* Suárez, A. M., Ramallo, F. & Rodríguez, X. P. (eds.), *Vigo, 23rd-26th October 2002*, Vigo, Servizo de Publicacións da Universidade de Vigo, pp. 305-318.

RHYS, M. & THOMAS, E. (2012), "Bilingual Welsh-English children's acquisition of vocabulary and reading: implications for bilingual education", *International Journal of Bilingual Education and Bilingualism*, 16, 6, 633-656.

RUIZ MORENO, M. (2010), *Spanish expressive vocabulary acquisition among Catalan bilinguals*, unpublished master thesis, Bangor, Bangor University.

RUIZ MORENO, M. (2020), *Phonetic production in early and late German-Spanish bilinguals*, PhD thesis, Hamburg, Hamburg University.

STEPENES, C. (2013), *Executive Function Development: A comparison of monolingual and bilingual children in Ireland*, PhD thesis, Belfast, Queen's University Belfast.

VAN RUIJVEN, E. (2006), *Leervorderingen in het Friese basisonderwijs: Onderzoek naar het onderwijsniveau en de leervorderingen van de leerlingen in het Friese basisonderwijs*, Leeuwarden, Fryske Akademy.