

DUAL LEXICAL ASSESSMENT OF BASQUE-SPANISH BILINGUALS USING THE MACARTHUR-BATES COMMUNICATIVE DEVELOPMENT INVENTORIES

MARIA-JOSÉ EZEIZABARRENA¹, IÑAKI GARCIA FERNANDEZ¹, AROA MURCIANO¹

¹University of the Basque Country (UPV/EHU), Spain

Corresponding author: Maria-José Ezeizabarrena. mj.ezeizabarrena@ehu.eus

Received: 31.10.2025

Accepted: 19.02.2026

ORIGINAL RESEARCH ARTICLE

UDK: 81'23:[811.134.2+811.361]

81'23:81'246.2

DOI: <https://doi.org/10.31299/hrri.62.si.5>

Abstract: This study focuses on the relationship between the amount of exposure to Basque and Spanish languages, and the expressive vocabulary of a group of 88 children (aged 16-30 months) living in the Basque Autonomous Community in Spain. The expressive vocabulary of these children in both languages and their relative amount of exposure were estimated using parental reports obtained with the standardised Basque and Iberian Spanish versions of the Communicative Development Inventory II (CDI II). A high degree of individual variability was found in the amount of exposure (more/less balanced relative input), as well as in vocabulary size (in one or both languages) of these early simultaneous bilingual children, even among those within the same age category (measured in months). Statistical analyses were performed after introducing age of participants as a covariate, and the results revealed a) a strong effect of the relative amount of exposure on the expressive vocabulary size in Basque, but no effect on the Spanish vocabulary size, and b) no effect on the conceptual vocabulary of the children. These findings are consistent with the claim that the amount of exposure affects language specific vocabulary size, more than conceptual vocabulary. The divergent results in the two languages indicate that the majority versus minority status of the language can influence lexical development in toddlers.

Keywords: bilingualism, Basque, Spanish, minority language, official language

INTRODUCTION

Children develop their language skills based on the language experiences they have. However, variability in contact situations, as well as in the ways in which bilingualism is operationalised across studies, makes it difficult to identify participants' profiles and categorise them as being strictly monolingual or a type of bilingual (Surrain & Luk, 2019). Thus, labels such as *simultaneous*, *sequential*, *(un)balanced bilingual*, or even bilingual *dominant* in language A or B, often vary depending on whether they are established based on age of acquisition, proficiency, language usage, (absolute/relative/cumulative) input, or the scores obtained in one task or another (Solís-Barroso & Stephanich, 2019).

In contrast with the widely-held belief that early simultaneous bilingualism is associated with balanced language exposure and parallel development in the two languages, many speakers

of minority languages are *early simultaneous* or *sequential unbalanced* bilinguals. This peculiarity makes the acquisition of minority languages a unique scenario for the study of language dominance, as well as for investigating the impact of the status (official/non-official, or majority/minority) of languages on the level of language accuracy in bilingual children (Gathercole & Thomas, 2009).

Basque is a non-Indoeuropean minority language spoken in several territories of Spain and France. For centuries, it has coexisted with at least two Romance languages, Spanish and French, and currently, it has a status that varies across territories (*regional* in France, *co-official* in some areas of Spain). A recent sociolinguistic survey conducted in the territory known as the Basque Autonomous Community (BAC), where Basque has been co-official with Spanish since 1979, reported that, in 2021, 45.3% of the population older than 15

years were monolingual speakers of Spanish, 0% were monolingual Basque speakers, while 54.8% were more or less active in the use of Basque (Eusko Jaurlaritza, 2024). The monolingual population decreases in younger age cohorts. Only 1.3% of 5- to 9-year-old children were monolingual speakers of Spanish, while most had high (80.9%) or low (17.7%) expressive competence in Basque (Eustat, 2021). These data suggest that Basque speakers are bilingual and regular users of Spanish in the BAC, and that the possibility of raising children in monolingual Basque environments continue to be limited, despite the revitalisation of Basque in recent decades in this territory.

In the absence of a monolingual population, the much-debated comparison between bilinguals and monolinguals does not apply to Basque. In line with the recommendation from Pearson and colleagues on assessing bilinguals in terms of both their languages, after taking the amount of exposure into account, the current study aims to investigate the strength of the relationship between the reported Basque and Spanish input estimates, and the size of the expressive vocabu-

lary in these languages (Pearson, 1998; Pearson et al., 1997). It also discusses the effect of language status (majority/minority) on the lexical development of young bilinguals.

METHOD

The present study analysed reported data on expressive vocabulary and the estimated amount of exposure to Basque (B) and Spanish (Sp) languages, based on a group of 88 children from B-speaking families. This data was collected using the Basque CDI and Spanish CDI questionnaires, both completed within a period of 7 days.

The cross-sectional sample included data collected from 88 toddlers between the ages of 16 and 30 months ($M = 23.11$ months). The sample was balanced in terms of gender (51.1% boys, 48.9% girls; See Table 1). All participants had some exposure to B and Sp, since they are being raised in families where at least one parent or tutor speaks to them in B, and they live in the Basque Autonomous Community, a region where the minority language, B, is co-official with Sp.

Table 1: Distribution of the sample according to age, gender, and input

	Age in months	Gender		Input			Total	
		Girl	Boy	Quasi-monolingual (> 90%)	B-dominant (61-90%)	Balanced (41-60%)		Sp-dominant (≤ 40%)
	16	1	1	1	0	1	0	2
	17	2	1	1	1	0	1	3
	18	7	1	2	2	4	0	8
	19	0	6	1	1	2	2	6
	20	2	6	3	4	1	0	8
	21	4	3	0	1	6	0	7
	22	2	3	1	1	3	0	5
	23	2	1	0	3	0	0	3
	24	8	3	2	8	0	1	11
	25	7	3	1	2	5	2	10
	26	4	3	2	3	1	1	7
	27	2	3	0	5	0	0	5
	28	2	4	0	1	5	0	6
	29	0	2	0	2	0	0	2
	30	2	3	0	1	0	4	5
	Total	45	43	14	35	28	11	88

The *relative amount of exposure* to B was calculated for each participant by converting into percentages the estimated number of hours per week that the informants reported observing family members speaking in B to the child, divided by the total number of hours of exposure to any language. Similar to previous Basque CDI studies, participants were categorised into four input groups: a) quasi-monolinguals (i.e.) children receiving > 90% relative exposure to B; b) B-dominant bilinguals, who receive between 61 and 90% B input; c) balanced bilinguals, who receive 41 to 60% of relative exposure to B, and d) Sp-dominant bilinguals, who receive ≤ 40% of relative B input (Ezeizabarrena et al., 2013).

Two standardised parental report instruments were used for the lexical assessment: the Basque (Barreña et al., 2008) and the Iberian Spanish versions (López-Ornat et al., 2005) of the Communicative Development Inventory: *Word & Sentences* (Fenson et al., 1993). Participants were asked to indicate which items their child produced from the corresponding checklists in the above-mentioned CDIs. All participants were asked to complete the questionnaires within a one-week interval. The Sp and the B checklists have a similar structure in terms of lexical classes, but the number of items varies between the Basque CDI (654 items) and the Spanish CDI (588 items). The instruments also differ in lexical items, since they share only 354 equivalent words (Examples 1a to 1c). Only a few B/Sp equivalent words are cognates (1c).

(1) B / Sp equivalent words and English translation

- a. *eskua / mano* ‘hand’
- b. *txakurra / perro* ‘dog’
- c. *azukrea / azúcar* ‘sugar’

Vocabulary was measured in words and concepts, following Pearson et al. (1997). The number of words marked in the standardised checklists were counted to determine each toddler’s expressive vocabulary size in B and Sp, respectively. The conceptual vocabulary was calculated by totalling the words produced in both languages, except for the equivalent pairs (Examples 1a to 1c), which were counted just once. To examine the influence of linguistic input on vocabulary size, univariate

analysis of variance (ANOVA) was performed in SPSS (version 28.0.1.1). The age of the subjects was introduced as a covariate. The data for the three variables analysed (words in B, words in Sp, and total concepts) did not follow a normal distribution (words in B: $Z(88) = 0.19, p < .001$; words in Sp: $Z(88) = 0.19, p < .001$; Total concepts: $Z(88) = 0.16, p < .001$). However, the tests for homogeneity of variance indicated equality of variances in the three variables analysed, according to Levene’s test (words in B: $F(3, 84) = 2.22, p = .09$; words in Sp: $F(3, 84) = 0.02, p = .99$; Total concepts: $F(3, 84) = 2.14, p = .10$). In these cases, univariate ANOVA is a robust test that adequately controls for Type I error (Blanca et al., 2017). When significant differences were found, pairs of groups were compared using the Bonferroni test.

RESULTS

In general, the average exposure was greater to B ($M = 65.2\%$, $SD = 20.6$) than to Sp ($M = 34.96\%$, $SD = 20.7$). Nevertheless, high levels of individual variation were observed in vocabulary size among the B-Sp bilinguals, across all age and language input groups (Fig. 1):

- Quasi-monolingual B children ($n = 14$, 15.9%), ranging in age from 16 to 26 months, were estimated to receive 91-100% relative B input and 0-9% complementary Sp input. Their vocabulary size was calculated as a total of 0 to 535 words in B ($M = 193.2$, $SD = 32.0$) and 0 to 499 words in Sp ($M = 112.7$, $SD = 26.0$). Although this may seem counterintuitive, two participants (ages 17 and 18 months, respectively) who were reported to be growing up in families where all members spoke to them exclusively in B, and were accordingly coded as having 0% Sp input, were not excluded from the sample of early B-Sp bilinguals. The decision was based on the assumption that Sp is present in the daily life of all current B speakers living in the area studied, and that child-directed speech is not the only component of language exposure, as confirmed by the varied expressive vocabu-

lary of the 18-month-old in both languages (15 Sp words, 12 lexical B-Sp equivalents, in addition to 27 B words, resulting in a total of 30 concepts).

- B-dominant bilinguals ($n = 35$, 39.8%), ranging in age from 17 to 30 months, were estimated to receive 61-90% relative B input. Their vocabulary size was calculated as 6 to 509 words in B ($M = 154.4$, $SD = 19.9$) and 3 to 347 words in Sp (3-347 w, $M = 75.5$, $SD = 16.2$).
- Balanced bilinguals ($n = 28$ children, 31.8%), ranging in age from 16 to 28 months, were estimated to receive 42-60% relative B input. Their vocabulary size was calculated as 3 to 546 words in B ($M = 130.7$, $SD = 22.2$) and 1 to 494 words in Sp ($M = 110.9$, $SD = 18.0$).
- Sp-dominant bilinguals ($n = 11$, 12.5%), ranging in age from 17 to 30 months, were

estimated to receive 17-40% relative B input. Their vocabulary size was calculated as 4 to 179 words in B ($M = 33.9$, $SD = 35.8$) and 10 to 481 words in Sp ($M = 143.6$, $SD = 29.1$).

Statistical analyses revealed a significant medium to large effect of relative input on B vocabulary across the entire sample ($F(3, 88) = 4.02$; $p = .01$; $\eta_p^2 = .13$). Post-hoc comparisons using Bonferroni's test revealed that only the Sp-dominant group (B input < 40%) differed from the other three input groups in B vocabulary size ($p < .05$; See Fig. 1). In contrast, no differences were found between input groups in Sp vocabulary size ($F(3, 88) = 1.75$; $p = .16$; $\eta_p^2 = .06$), nor in conceptual vocabulary ($F(3, 88) = 0.82$; $p = .49$; $\eta_p^2 = .03$). Interestingly, balanced bilinguals did not differ from B-dominants or from quasi-monolinguals ($p > .05$) in their B and Sp vocabulary, or in conceptual vocabulary.

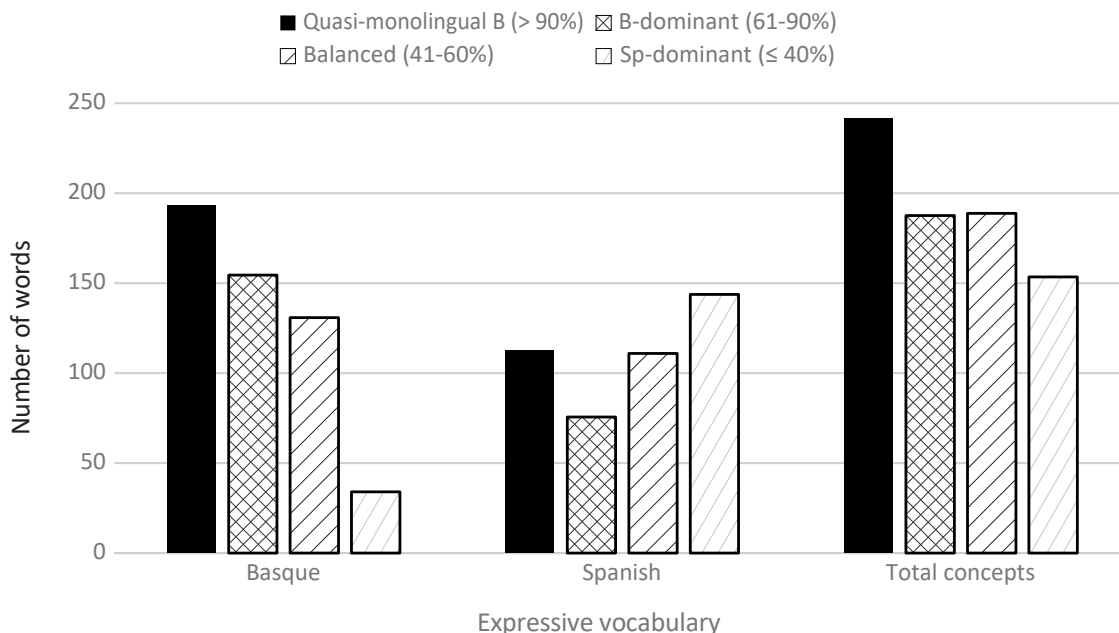


Figure 1. Estimated means of expressive and conceptual vocabulary size in four input groups, with age as covariate

DISCUSSION AND CONCLUSIONS

The variety of patterns found in the sample analysed highlights the mismatch between bilingual profiles based on age of acquisition, relative in-

put, and lexical development. In this context, the label *early* and/or *simultaneous*, which is based on age and order of acquisition, applies to all Basque-Spanish *bilingual* toddlers participating in the present study. However, it covers a spec-

trum of bilingual types, ranging from *B-dominant* to *balanced* and to *Sp-dominant bilinguals*, whose boundaries vary depending on whether they are based, for instance, on the amount of exposure or language knowledge (expressive vocabulary) of bilingual children. For instance, there were some cases in which children with balanced input had unbalanced expressive vocabulary sizes, and vice versa. In the present sample, an 18-month-old child with reported 100% exposure to B in the home (identified as monolingual B) but who had an expressive vocabulary in both in B and Sp languages could be categorised as either monolingual B or B-dominant bilingual. This case is an illustrative example of the inconsistencies and challenges that bilingual research faces, more specifically in the context of minority languages. The above-mentioned data could call into question the reliability of reported data in general, or of this specific sample in particular. However, the findings obtained with these two standardised instruments, the Basque and Spanish CDIs, are consistent with the results of a previous study conducted with a sample of B-Sp children including infants and toddlers (Ezeizabarrena et al., 2018).

Despite the individual variability mentioned, data have confirmed the strength of the association between input estimates and vocabulary learned in B, similar to the seminal paper by Pearson et al. (1997). Differences were found between input groups in expressive vocabulary in B, although not among all of them, in line with Thordardottir (2019) and Cattani et al. (2024). On the one hand, the absence of differences between balanced, B-dominant, and quasi-monolingual B input groups confirmed previous findings obtained with larger Basque CDI samples (Ezeizabarrena et al., 2013; Garcia et al., 2024). On the other, the significantly smaller B vocabulary size in the Sp-dominant group aligned with the threshold of 40% of relative input observed by Thordardottir (2019:253): “*children that received less exposure to French than English (39% or less) are the ones who consistently scored significantly lower than monolinguals in French [...] for receptive and*

expressive vocabulary”. See also Cattani et al. (2014), for a different cut-off point (60%) showed that the performance of bilingual toddlers was similar to that of monolinguals. The discussion about the similarities and differences between these pieces of research goes beyond the scope of the present study.

The strength of the effect of input on B vocabulary size and the lack of effect on Sp vocabulary shows the asymmetrical influence of the majority (Sp) versus minority language (B), even by the age of 2;6. This has been observed for English-Sp bilingual toddlers (Pearson et al., 1997) and for Welsh-English bilingual school-aged children, as stated by Gathercole & Thomas: “*In bilingual communities in which one language is very dominant, the acquisition of the dominant language may be unproblematic across bilingual sub-groups, while acquisition of the minority language can be hampered under conditions of reduced input*” (Gathercole & Thomas, 2009: 213). Finally, the lack of input effects on conceptual vocabulary is consistent with previous CDI studies (Pearson et al., 1997). See also Core et al. (2013).

All in all, these findings confirm the need to use the terms *language dominance/balance* in early language acquisition with caution, by taking exposure, us(age), and competence into account, without disregarding the status of the languages involved (Eckhaus & Ezeizabarrena, in press; Meisel, 2011; Montrul, 2016; Surrain & Luk, 2019; Thordardottir 2019).

Acknowledgement

We are very grateful to the University of the Basque Country (UPV/EHU), to the research groups KGNZ and ELEBILAB, and especially to all the parents and the staff of kindergartens and schools who made this study possible. This work was supported by grants from the Basque Government (IT983-16, IT1627-22; IT2040-26), MICIU/AEI/10.13039/501100011033 and FEDER UE (PID2023-148030NB-I00).

REFERENCES

- Barreña, A., García, I., Ezeizabarrena, M. J., Almgren, M., Arratibel, N., Olano, I. Barnes, J., Petuya, A., & Colina, A. (2008). *MacArthur-Bates komunikazio garapena neurtzeko zerrenda euskarara egokitua. Erabiltzaileentzako gida eta eskuliburu teknikoa*. Udako Euskal Unibertsitatea.
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option?. *Psicothema*, 29(4), 552-557. <https://doi.org/10.7334/psicothema2016.383>
- Cattani, A., Abbot-Smith, K., Farag, R., Krott, A., Arreckx, F., Dennis, I., & Floccia, C. (2014). How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests?. *International Journal of Language and Communication Disorders*, 49(6), 649-671. <https://doi.org/10.1111/1460-6984.12082>
- Core, C., Hoff, E., Rumiche, R., & Señor, M. (2013). Total and Conceptual Vocabulary in Spanish-English Bilinguals from 22 to 30 Months: Implications for Assessment. *Journal of Speech, Language and Hearing Research*, 56(5), 1637–1649. [https://doi.org/10.1044/1092-4388\(2013/11-0044](https://doi.org/10.1044/1092-4388(2013/11-0044)
- Eckhaus, R. & Ezeizabarrena, M.J. (in press). Dynamicity in the language dominance of a simultaneous bilingual preschooler with two home languages (English-Spanish) and a third societal language. In Pérez-Tattam, R.; Stadthagen, H.; Kaščelan, Dr. & Parafita-Couto, M.C. (eds.) *Current Insights into the Interaction of Linguistic and Extralinguistic Factors in Language Learning. Current Issues in Bilingualism*. Language Science Press.
- Eustat (2021). *Population aged 5 and over of the Basque Country by province and age group according to language skills. 1981-2021*. Available online on 23.12.2025: https://www.eustat.eus/elementos/tbl0009563_c.html
- Ezeizabarrena, M.J.; Barnes, J.; García, I.; Barreña, A. & Almgren, M. (2013). Using Parental Report Assessment for Bilingual Preschoolers: the Basque experience. In V. C. Mueller Gathercole (Ed.), *Solutions for the Assessment of bilinguals* (pp. 57-80), Multilingual Matters.
- Ezeizabarrena, M.J.; García, I.; Almgren, M. & Barreña, A. (2018). Diferencias individuales en el desarrollo léxico de bilingües simultáneos tempranos euskera-castellano. In C. Rojas & V. Oropeza (Eds.) *Diferencias individuales en la adquisición del lenguaje* (pp. 195-218). Universidad Autónoma de México.
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. Singular Publishing Group.
- García, I.; Ezeizabarrena, M.J. & Murciano, A. (2024). The Basque version of the CDI-Words and Gestures, extended up to age 2. *Applied Psycholinguistics*, 45, 624-646. <https://doi.org/10.1017/S0142716424000201>
- Gathercole, V.C.M. & Thomas, E.M. (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>
- López-Ornat, S., Gallego, C., Gallo, P., Karousou, S., Mariscal, S., & Martínez, M. (2005). *Inventario de desarrollo comunicativo MacArthur*. TEA ediciones.
- Meisel, J.M. (2011). *First and Second Language Acquisition*. Cambridge University Press.
- Montrul, S. (2016). Dominance and proficiency in early and late bilingualism. In C. Silva-Corvalan & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals* (pp. 15–35). Cambridge University Press.
- Pearson, B.Z. (1998). Assessing Lexical Development in Bilingual Babies and Toddlers. *The International Journal of Bilingualism*, 2 (3), 341-372.
- Pearson, B.Z., Fernandez, S.C., Lewedeg, V., & Oller, D.K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18, 41-58.
- Solis-Barroso, C. & Stephanich, S. (2019). Measuring language dominance in early Spanish/English bilinguals. *Languages* 4(3), 62. <https://doi.org/10.3390/languages4030062>

Surrain, S. & Luk, G. (2019). Describing bilinguals: a systematic review of labels and descriptions used in the literature between 2005-2015. *Bilingualism: Language and Cognition*, 22(2), 401-415. <https://doi.org/10.1017/S1366728917000682>

Thordadottir, E. (2019). Amount trumps timing in bilingual vocabulary acquisition: Effects of input in simultaneous and sequential school-age bilinguals. *International Journal of Bilingualism*, 23(1), 236-255. <https://doi.org/10.1177/1367006917722418>