Journal of Child Language (2025), 1–25
doi:10.1017/S0305000924000667

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Bilingual Vocabulary Development in Mexican Indigenous Infants: The Effects of Language Exposure from Home and Mothers' Language Dominance

Desarrollo del vocabulario bilingüe en infantes indígenas mexicanos: los efectos de la exposición a la lengua desde el hogar y de la dominancia lingüística de las madres

Stanislav Mulík[1,2] and Natalia Arias-Trejo[2]

[1]Department of Spanish, Italian, and Portuguese, The Pennsylvania State University, University Park, PA, USA and [2]Psycholinguistics Lab, Faculty of Psychology, National Autonomous University of Mexico, Mexico City, Mexico
Corresponding author: Stanislav Mulík; Emails: sxm6739@psu.edu; stanmulik@gmail.com

(Received 30 October 2023; revised 24 August 2024; accepted 19 November 2024)

Abstract
This study evaluates how language exposure and mothers' language dominance relate to infants' early bilingual vocabulary development in a low-socioeconomic status (SES) sample from an understudied population: Mexican Indigenous bilinguals. Thirty-two mother–child dyads participated. All mothers were bilingual speakers of Spanish and one of Mexican Indigenous languages, including Zapotec, Mixtec, and Otomi. Infants' (between 16 and 37 months) vocabulary size was estimated in both languages using the Mexican Spanish version of the MacArthur-Bates CDI II. Infants' language exposure, mothers' bilingual profile, and their SES were estimated on numerical scales. The results of Spearman correlations showed infants' vocabulary size in Spanish grows with age, while their vocabulary in the Indigenous language depends on relative language exposure. Mothers' language dominance correlated with Indigenous language exposure and infants' vocabulary size in the Indigenous language. These findings are discussed in the context of early bilingual vocabulary acquisition in speakers of minority languages.

Keywords: early language development; bilingual vocabulary; Indigenous bilingualism; language contact; minority language

Resumen

Este estudio evalúa cómo la exposición a la lengua y la dominancia lingüística de las madres se relacionan con el desarrollo temprano del vocabulario bilingüe de los/las bebés en una muestra de nivel socioeconómico bajo de una población poco estudiada: bilingües indígenas mexicanos. Treinta y dos díadas madre-hijo/hija participaron en el estudio. Todas las madres eran hablantes bilingües de español y una de las lenguas indígenas mexicanas, incluyendo zapoteco, mixteco y otomí. Los/las bebés tenían entre 16 y 37 meses de edad y se estimó el tamaño de su vocabulario en ambas lenguas utilizando la versión en español mexicano del CDI II de MacArthur-Bates. La exposición lingüística de los/las bebés desde el hogar, el perfil bilingüe de las madres y su estatus socioeconómico se estimaron en escalas numéricas. Los resultados de las correlaciones de Spearman mostraron que el vocabulario de los/las bebés en español crece con la edad, mientras que su vocabulario en la lengua indígena depende de su exposición relativa a esta lengua. La dominancia lingüística de las madres se correlacionó con la exposición a la lengua indígena y el tamaño del vocabulario de los/las bebés en la lengua indígena. Estos resultados se discuten en el contexto de la adquisición temprana de vocabulario bilingüe en hablantes de lenguas minoritarias.

Palabras clave: desarrollo temprano del lenguaje; vocabulario bilingüe; bilingüismo indígena; contacto lingüístico; lengua minoritaria

1. Introduction

Understanding early vocabulary development in infants is of great importance because, in modern societies, children's productive vocabulary size predicts their literacy competence (Lee, 2011), school success (Snowling et al., 2016), and even academic achievements later in life (Bleses et al., 2016). Infants usually utter their first words at around the age of 1, and at the age of 2, they already have about 100 words in their productive vocabulary (Jackson-Maldonado et al., 1993). However, the number of words 2-year-old children produce can vary greatly between 0 and more than 600 (Frank et al., 2017; Samuelson, 2021). The basic milestones of vocabulary development are in place for infants who are learning only one language (monolinguals) and also for those who are exposed to more than one language from birth (bilinguals), despite the fact that the latter have to split their time between two (or more) languages (Hoff et al., 2012; Petitto et al., 2001).

The problem is that our current understanding of early bilingual childhood development is mostly based on data from Western, educated, industrialised, rich, and democratic (WEIRD) societies (Henrich et al., 2010) with a bias towards higher socioeconomic status (SES). Participant recruitment from such samples hinders generalisability of results across world's populations (Arnett, 2008; Falk et al., 2013). Moreover, communication development in endangered languages, such as Mexican Indigenous languages, which are acquired in a Spanish-dominant context, might unfold differently (e.g., slower) than that of vital languages, such as English, Spanish, or French (Reese et al., 2018). The reason for these potential differences is not that the underlying learning mechanisms would differ for endangered languages. Rather, it is a matter of the social situation of Mexican Indigenous languages and their speakers, which often involves language endangerment, different levels of linguistic prestige, speaker discrimination, lack of educational language materials, and so on, but also living in poverty.

To the best of our knowledge, not much is currently known about early vocabulary development in bilingual Indigenous children, despite the fact that there are researchers (e.g., Pye, 2021) who call for reorienting the field towards studying the acquisition of Indigenous languages, many of which are rapidly disappearing. Including children from low-SES families and exposed to an endangered minority language and a vital majority language in studies on early bilingual vocabulary development is necessary in order for the field to move towards more inclusive and equitable research and to get a bigger picture of the issues that may arise in the process. Therefore, the present study provides novel data on early bilingual vocabulary development from a non-WEIRD and understudied population: Mexican Indigenous infants from low-SES contexts.

1.1. Early bilingual vocabulary development

Based on evidence from industrialised countries, bilingual infants develop language in a similar way to monolingual infants (Bialystok, 2001; Marchman et al., 2004; Parra et al., 2011). Importantly, bilingual infants are comparable to their monolingual peers with similar SES in total vocabulary size and acquisition rate (Hoff et al., 2012; Patterson, 2004; Patterson & Pearson, 2004; Pearson et al., 1993; Pearson & Fernández, 1994; Thordardottir, 2011). However, an important lag compared to monolingual children can present itself in vocabulary size of bilinguals when only one of the languages to which the child is exposed is evaluated (Gatt & O'Toole, 2016). Therefore, both languages of a bilingual infant are to be considered when assessing their total vocabulary size and to prevent misdiagnosing them with language development disorders when, in fact, they may just be developing in a way that is normal for bilinguals (Hoff, 2021).

Parental reports, such as the MacArthur–Bates Communicative Development Inventories (CDI; Fenson et al., 2012), are often used in studies of early bilingual vocabulary development to estimate infants' vocabulary size. Different language versions of CDI have been successfully implemented in monolingual and bilingual settings to provide a comprehensive inventory of infants' receptive and productive vocabulary (DeAnda et al., 2016), since young children typically understand more words (receptive vocabulary) than they can utter (productive vocabulary). In previous studies, bilingual infants' vocabulary size has been related to two main factors: relative language exposure and SES (DeAnda et al., 2016; Hoff, 2013). These two factors have been used in research as a proxy for linguistic input quantity and quality, respectively (De Cat, 2021), and are discussed in more depth in the following subsections.

1.2. Language exposure

Relative language exposure is directly linked to productive vocabulary size in bilingual infants (David & Wei, 2008; Pearson et al., 1997; Poulin-Dubois et al., 2013) as well as in slightly older children (Bedore et al., 2012): the more exposure to one of their languages the children have, the more words in that language they will say. Similar to its effects on vocabulary size, relative language exposure also correlates with the rate of vocabulary acquisition in bilingual children: the more vocabulary they already have, the faster they acquire new words (Hammer et al., 2008; Pearson, 2007; Pearson et al., 1997; Thordardottir et al., 2006). This is in line with studies carried out in monolingual children (Hoff, 2006).

Studies on the effects of language exposure on early vocabulary development are normally conducted in bilingual children who are exposed to well-known languages, such

as English, Spanish, and French. While these languages are vital majority languages spoken in various countries, they might also be minority languages in some bilingual settings (e.g., Spanish in the United States). Importantly, minority languages are acquired less easily than majority languages due mainly to the lack of environmental support (Hoff, 2021); in this sense, more relative exposure is needed for a minority language to be mastered to the same degree as the majority language (Pearson et al., 1997; Vihman et al., 2006). For example, in bilingual children who are learning one prestigious societal language (e.g., English) and one minority and/or endangered language (e.g., Irish), the relationships between language exposure and expressive vocabulary size might not follow the same pattern for both languages (O'Tooley & Hickey, 2019). In their study of bilingual Irish–English children of 17–36 months of age, O'Tooley and Hickey (2019) showed that there was a clear positive relationship between the exposure to English and the children's English vocabulary size. However, children's home exposure to Irish, their dominant language, did not clearly relate to their Irish vocabulary size unless English was not used at home at all. The importance of minority language exposure from home has recently also been investigated by Ezeizabarrena and García Fernández (2023), who studied bilingual children's vocabulary size in the minority Basque language in contact with majority languages, such as Spanish or French. The authors measured language exposure by the means of three exposure-related variables: relative input (<60%, 60%–90%, >90%), language of parental communication (speaking Basque either always, sometimes, or never), and parents' linguistic profile (one or both Basque-speaking parents). In their study, all three variables of language exposure to the minority language from home were shown to be significantly positively correlated with the minority language vocabulary size, with the effects becoming apparent between 18 and 29 months of age and growing in size during the following 2 years of the infant's life.

Based on this evidence, predominant relative exposure to the minority language from home seems to be a key factor in early bilingual vocabulary development, because children whose parents speak a minority language to them usually do not get additional exposure to this language outside the home. At the same time, however, exposure from home *only* might not be enough for children to acquire the minority language in the long run (Mancilla-Martinez & Vagh, 2013; Miękisz et al., 2016). In contrast, the majority language does not seem to even need to be spoken at home in order for the children to achieve a native-like competence in it later on, since they are exposed to it everywhere else (Duursma et al., 2007). This illustrates the imbalance of the effects of exposure on vocabulary size in early bilingual acquisition of minority and majority languages, but also the many different outcomes that may arise from different bilingual settings. Therefore, new language pairs as well as underrepresented bilingual populations are more than necessary to further investigate the effects of minority language exposure from home on children's productive vocabulary size. For this, the timeframe of 18–36 months seems to be especially informative.

1.3. Socioeconomic status

Low SES has been shown to negatively affect both children and adults in terms of their cognitive functioning, executive functions, and memory (Farah, 2017; Tooley et al., 2021). These, in turn, have adverse implications for many aspects of language acquisition, including early vocabulary development. For example, the vocabulary size of children

from typical American families at 36 months was found to be twice the size in high-SES contexts in comparison to low-SES contexts (Hart & Risley, 1995). Low SES has also been linked to smaller vocabulary sizes in older children (from 36 months to 13 years of age) from monolingual settings, independently of their race (Farkas & Beron, 2004). These effects of SES on language may be a result of the underlying effects of SES on children's brains. For example, children from low-SES have been reported to have lower functional development of their prefrontal cortex than children from mid/high-SES homes (Moriguchi & Shinohara, 2019).

As for children growing up with more than one language, early vocabulary development has been mostly reported in bilingual populations biased towards medium and high SES, while fewer studies have focused on low-SES bilingual samples (e.g., Fernald et al., 2013; Hoff, 2013; Hurtado et al., 2008). In the latter, low SES has been linked to smaller vocabulary sizes in children from multilingual settings (Fernald et al., 2013; Singh et al., 2023). Importantly, however, low SES is commonly conflated with bilingualism in speakers of minority languages, such as Spanish-speaking children learning English in the United States (Hoff, 2013). In an effort to tease apart the effects of bilingualism and SES, Meir and Armon-Lotem (2017) reported that bilingualism can negatively affect children's vocabulary size in each language (when compared to monolinguals) while not hindering their cognitive abilities, whereas low SES clearly shows negative effects on both children's linguistic abilities and their cognitive development. In other words, in low-SES speakers of minority languages, the problem for language development is not their bilingualism as such, but the situation of poverty they are immersed in.

Before we discuss these SES effects any further, please bear in mind that SES is an outcome of systemic issues that place families and their children at risk, rather than the parents themselves being at fault for not doing enough for their children. That said, the effects of SES on children's vocabulary size have been shown to be mediated via parents' education and book reading practices (Singh et al., 2023), as well as by the differences in child-directed speech complexity (Hoff, 2003). In addition to the differences in input quality, mothers from low-SES contexts also speak to their children less than mothers from high-SES contexts (Hoff, 2003), which might reduce the input quantity that children get in their mother tongue. There might be several other underlying causes for the effect of SES on vocabulary size, since SES is directly related not only to the economic dimension, but also to health, psychological well-being, educational attainment, and occupational status.

In the particular context of Mexico, children who grow up listening to an Indigenous language generally belong to lower SES than children who only grow up listening to Spanish. This is because the Indigenous people are among the poorest in Mexico: in 2018, 70% of the Indigenous population of Mexico was living in poverty, compared to 39% of the non-Indigenous (CONEVAL, 2019). This commonly translates to varying degrees of formal education (including little to no schooling), high levels of illiteracy, malnutrition, obesity, among others. Although Mexican Indigenous people constitute only less than 10% of the total population nowadays (CDI, 2016), about 85% of the Mexican Indigenous population is bilingual (INEGI, 2011). Therefore, bilingualism is a default reality of the majority of Mexican Indigenous population, and it has to be considered in Indigenous language studies. At the same time, truly low-SES samples are underrepresented in early bilingual vocabulary development research, so including poverty-stricken populations in such studies is imperative.

1.4. The present study

The present study builds on the current knowledge of early vocabulary development in monolingual and bilingual infants from WEIRD populations where CDIs are readily available for their cultures and languages (Dale & Penfold, 2011). As a result, early vocabulary development in these settings is relatively well-documented, including the effects of language exposure and SES on infants' vocabulary size (see Sections 1.2 and 1.3).

In sharp contrast to this, only a handful of studies have investigated bilingual Indigenous children's early vocabulary acquisition using CDIs. For example, Vogt et al. (2015) adapted the CDI I and II (short versions) into two languages of Mozambique, Changana and Ronga. However, when the CDI was applied to a bilingual Ronga-Portuguese situation, the researchers asked the parents to report whether their 12-to-25-month-old children produced the word either in Ronga or Portuguese, thus only measuring their total conceptual vocabulary and not reporting vocabulary size in each language separately. In a similar light, when developing an adapted version of CDI for Xhosa-speaking children in the 16-to-30-month age range, Dowling and Whitelaw (2018) acknowledged the difficulty of teasing apart between Xhosa and English vocabulary, since many of the children in the pilot study who were declared monolingual in Xhosa by their parents preferred the English words for various everyday objects, while using Xhosa words for others. Although integrating scores from two languages seems to be a valid and useful strategy with bilingual infants and toddlers (Mancilla-Martinez et al., 2011), it does not provide detailed information about vocabulary composition between their two languages.

As for measuring vocabulary in both languages separately, Reese et al. (2018) conducted a study in New Zealand, which reported on 24-month-old children's vocabulary sizes in Māori and English. They found an adverse effect of SES on children's total vocabulary size; however, the study design did not include the effect of language exposure on vocabulary size in either language. Therefore, what remains to be investigated is how parental input in the Indigenous language and the majority language influences Indigenous infants' early bilingual vocabulary development, since these effects might be different to those reported in the literature for well-described and vital languages.

To address these gaps in the literature, the present paper describes productive vocabulary development in terms of vocabulary size and bilingual composition, in a sample of Mexican Indigenous infants who are exposed to an Indigenous language from home as well as to Spanish. It is important to mention that most Mexican Indigenous languages are not related to each other, let alone to Spanish. In present-day Mexico, there are 11 Indigenous language families, which comprise 68 language groups with 364 individual language varieties (INALI, 2008). For example, Otomi speakers do not understand Zapotec, and vice versa. Therefore, speakers of different Indigenous languages usually speak Spanish in order to communicate with each other. What these different speakers do have in common is the context of Indigenous language use, which is mostly oral and in the home, as opposed to Spanish, which is the language that is both spoken and written, and used in most communicative situations in the Mexican society. While we are aware that linguistic properties, such as how typologically different the two languages of a bilingual are, may influence dual language learning (Floccia et al., 2018) and that differences in input quality may be sources of great variability in dual language learners outcomes (Place & Hoff, 2016), in the present work the rationale behind grouping the bilingual speakers of Spanish and one of eight different Mexican Indigenous languages is their status as minority language bilinguals who live in low-SES homes, as well as the commonalities in the context of use of the minority language by Indigenous language speakers.

Another important aspect of the present study to mention is that we used the Mexican Spanish version of MacArthur–Bates Communicative Development Inventory – CDI II (Jackson-Maldonado et al., 2003) to estimate children's productive vocabulary size in both of their languages. It was beyond the scope of this work to develop a translated *and* adapted version of the CDI II for each of the eight Mexican Indigenous languages considered here and, therefore, the Mexican Spanish word list was used to estimate the size of children's productive vocabulary in both of their languages. Crucially, this still allowed us to explore the relationships between infants' vocabulary size in both languages and their exposure to the Indigenous languages as well as the degree of their mothers' bilingualism in terms of relative language dominance.

Based on previous studies, we hypothesised that there would be a positive effect of relative language exposure on infants' vocabulary size in both of their languages, but more relative exposure to the minority (Indigenous) language might be needed for the infants to reach a comparable vocabulary size as in the majority language (Spanish). The latter prediction is related to the fact that the use of Mexican Indigenous languages is overall discriminated and stigmatised and that there is a lack of support for their acquisition outside the home. Also, we predicted smaller productive vocabulary sizes in children who are growing up in lower SES settings than those from higher SES. Finally, we expected the mothers' degree of bilingualism to be related to children's vocabulary size in both of their languages.

2. Method

2.1. Participants

A total of 32 Mexican Indigenous mother–child dyads participated in the study. All 32 mothers were bilingual to different degrees, but all were speakers of Mexican Spanish and one of the following Mexican Indigenous languages (the number of speakers in the sample is given in parentheses): Zapotec (14), Mixtec (6), Otomi (4), Mazatec (3), Nahuatl (2), Mixe (1), Chinantec (1), and Triqui (1). These mothers were included in the study because they confirmed that the Indigenous language was spoken in the home where their child was growing up, that the child's age was approximately between 18 and 36 months, and that the child had normal vision and hearing and no history of neurological disorders.

The children's ages ranged from 16 to 37 months of age (mean age 27 ± 7 months). The mothers' ages ranged from 23 to 42 years (mean age 31 ± 5 years). In terms of age of acquisition (AoA), 9 of the 32 mothers learned both languages from birth (AoA = 0), while 19 learned the Indigenous language from birth with Spanish as a second language (AoA range 3–15 years) and 4 learned Spanish from birth with the Indigenous language as a second language (AoA range 2–10 years). The average AoA of Spanish in the mother sample was 5 ± 5 years, while the average AoA of Indigenous language was 1 ± 2 years.

Participants were interviewed in three states of the Mexican Republic: Querétaro, Baja California, and Oaxaca (Figure 1). Within the first group of five participants in total, four were interviewed in Amealco de Bonfil, in the south of the state of Querétaro, and one in the capital of the state of Querétaro, Santiago de Querétaro. Eight participants who belong to the second group were migrants from the state of Oaxaca, interviewed in San Quintín, Baja California, where they lived at the time of the study. The third group consisted of 19 participants, 13 of whom lived in the capital of the state of Oaxaca, Oaxaca de Juárez, and 6 in surrounding communities (3 in San Pablo Güilá and 3 in San Miguel del Valle).

Figure 1. Data collection sites within the Mexican Republic. A: Querétaro; B: Baja California; C: Oaxaca.

2.2. Materials

The Mexican Spanish version of MacArthur–Bates Communicative Development Inventory – CDI II (Jackson-Maldonado et al., 2003) was used to estimate the size of children's productive vocabulary, both in Spanish and in the Indigenous language. At the present time, there are no versions of CDI II in the Mexican Indigenous languages that were spoken by the participants of this study. Using the Mexican Spanish version of CDI II, of course, implies several limitations to the vocabulary size estimation in the Indigenous languages. Although Indigenous communities are a part of Mexican culture, the version used might not be completely culturally appropriate for all Mexicans, since cultural and ethnic diversity of Mexico is not homogeneous. Second, the inventory used in this study is normally used up to 30 months of age and, herein, we considered children of up to 37 months of age. However, this instrument was previously used on Spanish–English bilingual children from low-SES families in the United States and was shown to successfully capture bilingual vocabulary development between 24 and 36 months of age, despite the 30-month-old ceiling in high- and medium-SES populations (Mancilla-Martinez et al., 2011). Finally, it is reasonable to assume that imprecisions would also arise from the task of the mothers to report on their children's vocabulary not in one, but in two languages at the same time. Considering these limitations, the CDI II was deemed by the authors of this paper the most appropriate instrument for the children's vocabulary size estimation in both languages, since there is presently no better way of doing so. This is also in line with Kelly et al. (2015), who call for being pragmatic in Indigenous language assessment and using the best available method at the time.

Similar to vocabulary size, relative language exposure is also usually reported in the form of a parental report, which has proven to be a reliable measure of language exposure (e.g., Bedore et al., 2012; Marchman et al., 2004; Parra et al., 2011; Thordardottir et al., 2006). In the present work, infants' language exposure and use were estimated by using a

simple questionnaire consisting of 11 questions and based on the parental report from Duursma et al. (2007). The survey enquires about the languages spoken in the household (both at the infant and by the infant), about who speaks them, and for how many hours per week are they spoken by the different people. The questions were the following:

1. What language does the child's mother speak to them?
2. What language does the child's father speak to them?
3. What language do other adults in the home (other than the parents) speak to the child?
4. What language do the other children in the home speak to the child?
5. What language does the child speak to their mother at home?
6. What language does the child speak to their father at home?
7. What language does the child speak to other adults in the home (other than the parents)?
8. What language does the child speak to their friends outside the home?
9. How many hours a day is the child with their mother?
10. How many hours a day is the child with their father?
11. How many hours a day is the child with other adults in the home?

To measure participants' SES, CNS AMAI 2022 (Asociación Mexicana de Agencias de Inteligencia de Mercado y Opinión AC, 2022) was used, which is an SES calculator developed and used in Mexico. It consists of six easy-to-answer questions to quickly determine the respondent's SES on a continuum of SES scores ranging from 0 to 300, while also providing the pertinence of the participant to a social class on a 7-level scale that reflects the SES division of the Mexican society. The questions cover the following topics: educational level of the head of household, number of complete bathrooms in the home, number of cars in the home (understood as the sum of cars, vans, and pickups in the home), Internet connection at home, number of household members over 14 years of age who work, and number of bedrooms in the home. Each answer comes with a score, which adds to form the final score.

To assess mothers' dual-language background and usage patterns, we administered the Bilingual Language Profile (BLP) questionnaire (Birdsong et al., 2012; Gertken et al., 2014), since it provides detailed information on relevant determinants of bilingualism, while being easy to use and quick to apply. Crucially, BLP was previously successfully used in Mexico with an Indigenous population (an Otomi-Spanish bilingual population) to describe their degree of bilingualism in these languages (Mulík et al., 2021a).

The BLP is divided into four modules, each contributing equally to the overall language score. These modules are language history, language use, language competence, and language attitudes. The language history module assesses the age at which the bilingual began learning their languages and became comfortable using them. It also considers the total number of years of formal classes in both languages, time spent in a region where each language is spoken, and use of each language with family, friends, colleagues, or classmates. The language use module asks bilinguals to report the percentage of time they use each language in a typical week, including communication with others, talking to themselves, and mental calculations. The language competence module collects information on participants' perceived competence in their two languages, including oral and written expression, listening comprehension, and reading ability. The language attitudes module of the BLP questionnaire asks about participants' attitudes

towards their two languages, such as whether they identify with the culture represented by each language, and whether it is important for them to be perceived as native speakers of each language.

Language dominance is calculated based on the data collected from the four modules. The BLP produces a partial score for each module and for each language. The total score for each language is calculated by weighting the values obtained in the four different modules (Birdsong et al., 2012). The global linguistic dominance score is calculated by subtracting the total score obtained for Spanish from the total score obtained for the Indigenous language. This way, negative numbers designate linguistic dominance towards Spanish, positive numbers suggest language dominance towards the Indigenous language, and global dominance scores close to zero identify a relative balance of the two languages in the bilingual. The higher the absolute value of the BLP score, the more monolingual the person; the closer to zero the absolute value, the more bilingual the person.

2.3. Procedure

The data were collected as part of a larger study on processes involved in novel word learning in Mexican Indigenous infants. Before it was carried out, all participating mothers were informed about the aims of the study, about the instruments that would be used, and about the procedures that would be implemented during and after the study. Once the participants were satisfied with how all of their questions were answered by the researcher, they gave their written consent to voluntarily participate in the study. The informed consent form, and all the procedures and instruments used in the study, were approved by the Research Ethics Committee of the Faculty of Psychology at the Autonomous National University of Mexico. In order to protect participants identity, the collected data were anonymised before analysis.

2.4. Data analyses

Productive vocabulary size of the infants was estimated both in Spanish and in the Indigenous language. The data were obtained from part A (Vocabulary Checklist) of CDI II by orally asking the mothers, word by word, whether their child said the word in the Indigenous language, in Spanish, in neither language, or in both languages. For 27 out of 32 children, all 680 items in the 23-word categories of the Vocabulary Checklist were covered. These children's vocabulary size was estimated by adding the number of words they said in each language. For the remaining five children, only a part of the Vocabulary Checklist was covered. These children's vocabulary size was estimated by calculating what was the proportion of vocabulary that the filled in part represented in the rest of the sample (in the 27 children) and then the sum of words was multiplied by this number for the remaining 5 children.[1] The number of translation equivalents was estimated for each child by counting the number of words they said in both languages. Total and conceptual productive vocabulary was calculated for each participant as well, with the total productive vocabulary being the sum of the words said in Spanish and the words said in the

[1]The statistical analyses were carried out in duplicate: once on the complete sample of 32 dyads, and once on a reduced sample of 27 dyads, leaving the five children with extrapolated data out. Crucially, both analyses yielded the same results. We report only the first analysis with the complete sample of 32 dyads.

Indigenous language, while the conceptual productive vocabulary being the total productive vocabulary minus the number of translation equivalents (Core et al., 2013; Pearson et al., 1993). Despite total and conceptual vocabulary being both a common measure of bilinguals' vocabulary in the literature, they may differ in several aspects (Byers-Heinlein et al., 2024). For example, total vocabulary of bilingual children between 14 and 36 months has been shown to be similar (Core et al., 2013; Hoff et al., 2012) or even higher (Bosch & Ramon-Casas, 2014; Junker & Stockman, 2002) when compared to monolinguals. On the other hand, the measure of conceptual vocabulary might also show similar measures between bilinguals when comparing them to monolinguals (De Houwer et al., 2014; Junker & Stockman, 2002) or show bilinguals to have lower scores than monolinguals (Core et al., 2013; Thordardottir et al., 2006). In the present work, we decided to use both measures of expressive vocabulary size to get a more comprehensive picture of children's abilities. Finally, we also calculated infants' vocabulary composition, expressed as the percentage of Indigenous language words in their vocabulary, which was estimated by calculating the proportion of vocabulary they said in the Indigenous language of the mother by dividing the number of productive words estimated in their Indigenous language by the total productive vocabulary estimation. This was used as a measure of balance between the two languages, to give us a better idea of which language the children produced more words in, relative to the other language.

The infants' language exposure was estimated from the answers given by their mothers in the questionnaire. Their language exposure to the Indigenous language was calculated as relative to their exposure to Spanish, together adding to 100%. It is important to mention that even those children whose mothers reported them having no exposure to an Indigenous language were included in the analyses. The reason for this is that this measure is an estimate of direct exposure, which does not account for the children possibly overhearing the language. This way, even children with reported 0% exposure to the Indigenous language can be reported to have some Indigenous words in their vocabulary (which was in fact the case for two children in our sample). On the other hand, there is also the possibility of parents not reporting on their children's exposure to the minority language consistently with reality, as has been reported before for Polish children in the United Kingdom and Ireland (Miękisz et al., 2016).

The SES data were evaluated according to the guidelines that the questionnaire provides. Each question is assigned a score, and the sum of all the points that the participants get in the question is the final score that is used to estimate the SES on the scale. The seven social classes of the Mexican society are reflected in the SES ranges as follows (social class [SES range]): very low [0–47], low [48–94], low-high [95–115], medium-low [116–140], medium [141–167], medium-high [168–201], and high [202–300].

Mothers' language dominance on a numerical scale was obtained from the BLP questionnaire and used for analyses. Mother's dominant language was determined by whether their total BLP score was negative (Spanish-dominant) or positive (Indigenous language-dominant). Please note the difference between small (e.g., 1) and high (e.g., 132) positive scores; while for both the dominant language is Indigenous, a BLP score of 132 out of 218 suggests that the mother's relative language dominance is very skewed towards the Indigenous language, whereas a BLP score of 1 is very close to zero and, therefore, represents an almost perfectly balanced bilingualism between the slightly more dominant Indigenous language and slightly less dominant Spanish. The same applies to negative BLP scores.

As for statistical analyses of the data, most of the studied variables showed data that were not normally distributed (in both Kolmogorov–Smirnov and Shapiro–Wilk tests).

Therefore, the nonparametric Spearman's correlations were used to analyse relationships between the studied variables. In terms of strength, those values of $r_{Spearman}$ that are closer to zero describe weaker associations, whereas the closer the $r_{Spearman}$ values are to 1 or −1, the stronger the association between the studied variables is. As for significance, p <0.050 is considered a significant result of the statistical test.

3. Results

3.1. Overall results

The estimates of productive vocabulary size of the 32 children in the sample, as measured by CDI II, ranged from 3 to 628 words in Spanish (mean vocabulary size 194 ± 190 words) and from 0 to 200 words in the Indigenous language (mean vocabulary size 32 ± 57 words) from a total of 680 word items on the list. The children's exposure to the Indigenous language from home ranged from 0% (100% exposure to Spanish) to 95% (5% exposure to Spanish), with the mean exposure to the Indigenous language in the sample as a whole being 29 ± 33%. The mothers' bilingualism was on the spectrum from being Spanish-dominant (−132 out of −218) to being dominant in the Indigenous language (132 out of 218), with some of the mothers also being balanced bilinguals (scores around 0); the mean value of the dominance score of the sample being −26 ± 62, making the sample Spanish-dominant on average. The SES scores ranged from 29 to 152 out of 300, the mean value being 94 ± 29 and the bulk of the data belonging to low SES. This information is displayed in Table 1 for each participant separately.

3.2. The effects of infants' age on their vocabulary size

Figure 2 shows the effects of infants' age on their estimated vocabulary size, regarding conceptual, total, Spanish, and Indigenous language vocabulary. The results of Spearman correlations revealed that infants' age is positively related to both conceptual ($r_{Spearman}(32) = 0.554$, $p = 0.001$) and total ($r_{Spearman}(32) = 0.558$, $p = 0.001$) vocabulary size. However, when separated by language, only Spanish vocabulary size ($r_{Spearman}(32) = 0.523$, $p = 0.002$) but not the Indigenous language vocabulary size ($r_{Spearman}(32) = 0.060$, $p = 0.745$) correlates with infants' age.

3.3. The effects of infants' language exposure on their vocabulary size and composition

Conversely to the effects of age on infants' vocabulary size, the results of partial Spearman correlations whilst controlling for age revealed that infants' relative exposure to the Indigenous language (and, therefore, also to their relative exposure to Spanish) is not significantly related to either conceptual ($r_{Spearman}(29) = -0.043$, $p = 0.818$) or total ($r_{Spearman}(29) = 0.020$, $p = 0.913$) vocabulary size. When separated by language whilst controlling for age, only the Indigenous language vocabulary size ($r_{Spearman}(29) = 0.657$, $p < 0.001$) but not Spanish vocabulary size ($r_{Spearman}(29) = -0.245$, $p = 0.183$) correlated with the amount of infants' exposure to the Indigenous language. Figure 3 shows the effects of infants' exposure to the Indigenous language on their vocabulary size in Spanish and in the Indigenous language.

In terms of vocabulary composition, the proportion of Indigenous vocabulary in infants' total vocabulary is related to the amount of exposure they have to the Indigenous

Table 1. Participants' characteristics. *Positive BLP values designate mothers dominant in the Indigenous language; negative BLP designate Spanish-dominant mothers; BLP values close to zero designate relatively balanced bilinguals. These are not categories, but a continuum of relative language dominance: the higher the absolute value of the BLP score, the more monolingual the mother; the closer to zero the absolute value, the more bilingual the mother

Participant ID	Sex	Age (months)	Spanish vocabulary size	Indigenous vocabulary size	Total vocabulary size	Conceptual vocabulary size	Exposure to Indigenous language (%)	Indigenous language	Mother's BLP*	Mother's dominant language	SES (0–300)
Pa34	F	16	35	0	35	35	10	Zapotec	−58	Spanish	106
Pa62	F	16	28	1	29	29	1	Zapotec	−109	Spanish	117
Pa02	M	17	16	0	16	16	18	Otomi	−72	Spanish	82
Pa63	F	18	6	0	6	6	1	Chinantec	−34	Spanish	105
Pa68	F	18	8	30	38	38	75	Zapotec	39	Zapotec	104
Pa05	F	19	109	19	128	113	78	Nahuatl	1	Nahuatl	90
Pa03	F	22	50	0	50	50	20	Otomi	1	Otomi	111
Pa33	F	22	52	3	55	53	50	Zapotec	17	Zapotec	118
Pa37	F	22	348	1	349	348	0	Zapotec	−71	Spanish	78
Pa69	F	22	31	159	190	187	50	Zapotec	9	Zapotec	52
Pa60	M	23	155	69	224	159	20	Nahuatl	67	Nahuatl	66
Pa66	F	23	3	31	34	31	95	Zapotec	35	Zapotec	111
Pa65	F	24	120	1	121	121	80	Zapotec	29	Zapotec	64
Pa30	F	26	48	2	50	50	40	Zapotec	−62	Spanish	146
Pa35	F	26	397	117	514	403	50	Mazatec	−31	Spanish	146
Pa25	F	26	180	0	180	180	0	Zapotec	−35	Spanish	69
Pa27	M	28	419	0	419	419	20	Zapotec	−33	Spanish	64
Pa21	M	28	163	1	164	164	4	Triqui	33	Triqui	50

Table 1. *(Continued)*

Participant ID	Sex	Age (months)	Spanish vocabulary size	Indigenous vocabulary size	Total vocabulary size	Conceptual vocabulary size	Exposure to Indigenous language (%)	Indigenous language	Mother's BLP*	Mother's dominant language	SES (0–300)
Pa67	F	29	244	79	323	315	90	Zapotec	20	Zapotec	97
Pa08	M	30	366	0	366	366	2	Mixtec	−72	Spanish	116
Pa12	M	31	270	150	420	290	20	Mixtec	26	Mixtec	119
Pa32	F	31	511	0	511	511	10	Zapotec	−91	Spanish	94
Pa20	M	32	74	0	74	74	0	Mixtec	−116	Spanish	29
Pa31	M	32	118	0	118	118	0	Mazatec	−132	Spanish	83
Pa36	M	33	478	0	478	478	20	Mazatec	−111	Spanish	72
Pa06	M	35	400	0	400	400	10	Mixtec	−48	Spanish	152
Pa61	F	35	11	2	13	11	40	Mixtec	5	Mixtec	89
Pa64	M	35	11	140	151	144	95	Zapotec	132	Zapotec	119
Pa01	M	36	628	0	628	628	0	Otomi	−30	Spanish	108
Pa04	M	36	582	4	586	582	4	Otomi	−132	Spanish	74
Pa13	F	36	270	200	470	330	60	Mixtec	25	Mixtec	87
Pa18	F	37	70	10	80	80	0	Mixe	−45	Spanish	78

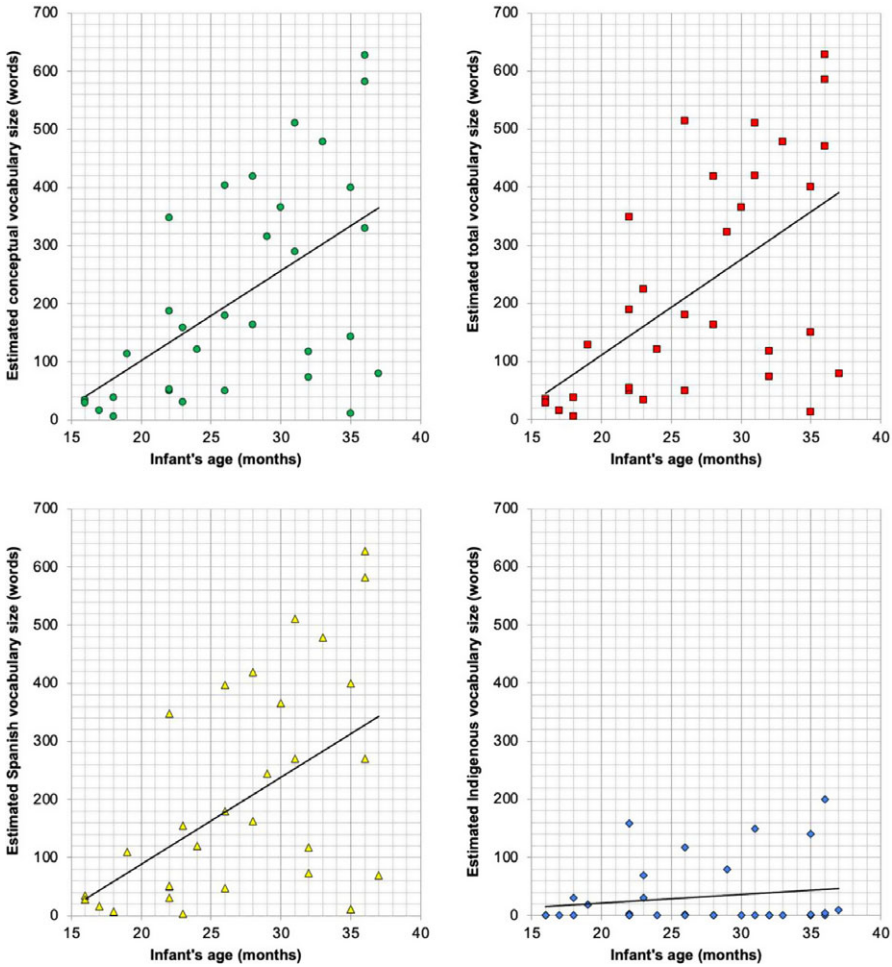

Figure 2. The effects of infants' age on their vocabulary size.

language ($r_{Spearman}(29) = 0.705$, $p < 0.001$) whilst controlling for age. Figure 4 illustrates this relationship. It is evident from Figure 4 that this is not a completely proportional, one-to-one relationship, because 50% of language exposure does not result in 50% of vocabulary proportion, but it is rather skewed towards the majority language (Spanish). Thus, according to the data obtained in our sample, in order to achieve 50% of vocabulary in each language of the bilingual infants, as much as 82% of exposure to the Indigenous language (and only 18% of relative exposure to Spanish) is needed (Figure 4).

3.4. The effects of socioeconomic status on infants' vocabulary size

No effect of SES on infants' vocabulary was found (data plots not shown). The results of partial Spearman correlations whilst controlling for age revealed that the SES of infants' home is not significantly related to either conceptual ($r_{Spearman}(29) = -0.103$, $p = 0.583$)

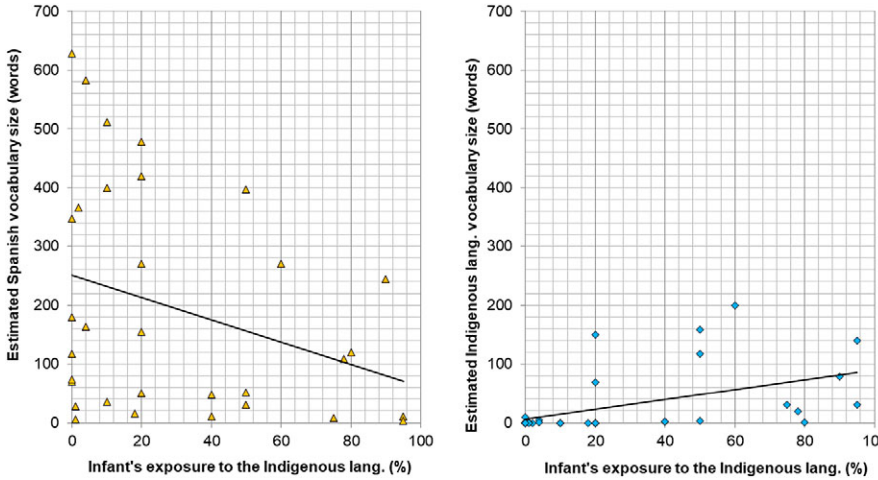

Figure 3. The effects of infants' exposure to the Indigenous language on their vocabulary size in Spanish (*left*, nonsignificant) and in the Indigenous language (*right*, significant).

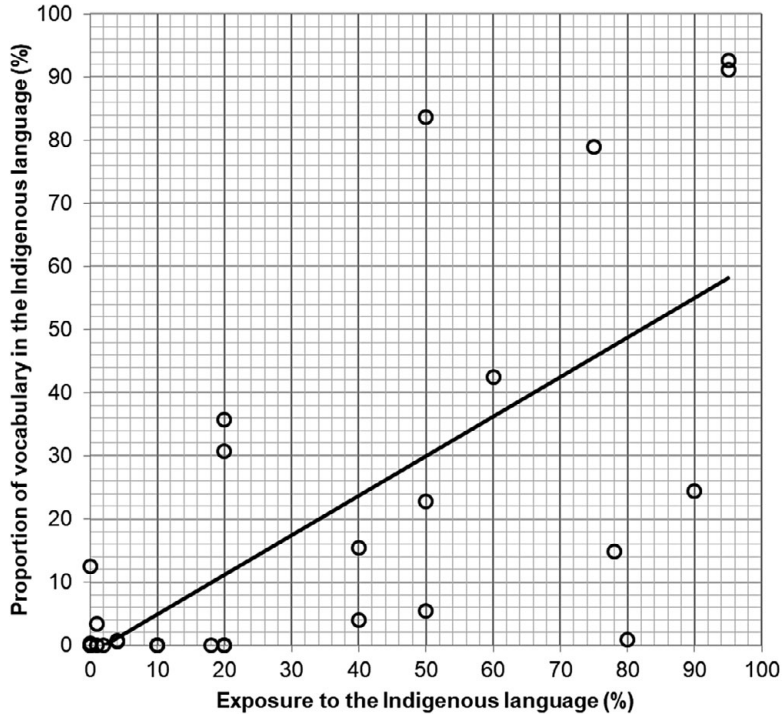

Figure 4. The effects of infants' exposure to the Indigenous language (%) on the proportion of their vocabulary in the Indigenous language (%).

or total ($r_{Spearman}(29) = -0.050$, $p = 0.790$) vocabulary size. Similarly, when separated by language whilst controlling for age, neither the Indigenous language vocabulary size ($r_{Spearman}(29) = 0.129$, $p = 0.488$) nor the Spanish vocabulary size ($r_{Spearman}(29) = -0.114$, $p = 0.542$) correlates with the infants' SES.

3.5. The effects of mothers' bilingualism on infants' language exposure and vocabulary size

Figure 5 shows the effects of mothers' bilingual profile on infants' exposure to the Indigenous language, the proportion of Indigenous language vocabulary, and the infants' estimated vocabulary size in both languages. As for exposure to the Indigenous language, a positive partial correlation whilst controlling for infants' age was found between this variable and the mothers' bilingual profile ($r_{Spearman}(29) = 0.671$, $p < 0.001$), in a sense that the more dominant the mothers are in the Indigenous language, the more exposure to the Indigenous language the infants get (Figure 5a). However, while Spanish-dominant mothers (negative BLP scores) are raising children with a relatively low proportion of Indigenous words in their productive vocabulary, there is a much greater variability in the proportion of Indigenous vocabulary for infants whose mothers are relatively balanced bilinguals (BLP scores around 0) or Indigenous language-dominant bilinguals (positive BLP scores) (Figure 5b). Finally, the mothers' language dominance partially correlates with vocabulary size in the infants' Indigenous language whilst controlling for their age (Figure 5c). Infants' Indigenous vocabulary size positively correlates with the mothers' BLP scores ($r_{Spearman}(29) = 0.645$, $p < 0.001$), while for their Spanish vocabulary size, there is a nonsignificant trend of negatively correlating with the mothers' BLP scores ($r_{Spearman}(29) = -0.296$, $p = 0.106$). Although the correlation is only significant for the Indigenous language and not for Spanish, it is evident from Figure 5c that these relationships are skewed in favour of Spanish.

4. Discussion

The aim of this study was to evaluate how Mexican Indigenous bilingual infants' exposure to the Indigenous language predicts their vocabulary development both in the Indigenous language and in Spanish, the dominant language of Mexican society. The present study not only adds to the growing body of research that deals with early bilingual vocabulary development in general, but especially it also provides novel data relevant to low-SES samples from understudied bilingual populations, such as speakers of endangered Indigenous languages. The key finding of our study is that, whilst controlling for infants' age, their relative language exposure satisfactorily explains only their expressive vocabulary size in the Indigenous language, but not the majority Spanish language. Moreover, the mothers' relative language dominance seems to be related to children's vocabulary size in both languages, although the association was only statistically significant for the Indigenous language. These results suggest that the higher the mother's relative language dominance towards the Indigenous language is, the larger the infant's productive vocabulary size in the Indigenous language will be. However, this also means that the higher the mother's relative language dominance towards Spanish, the smaller the infant's productive vocabulary size in the Indigenous language. These results are discussed in greater detail in the following paragraphs.

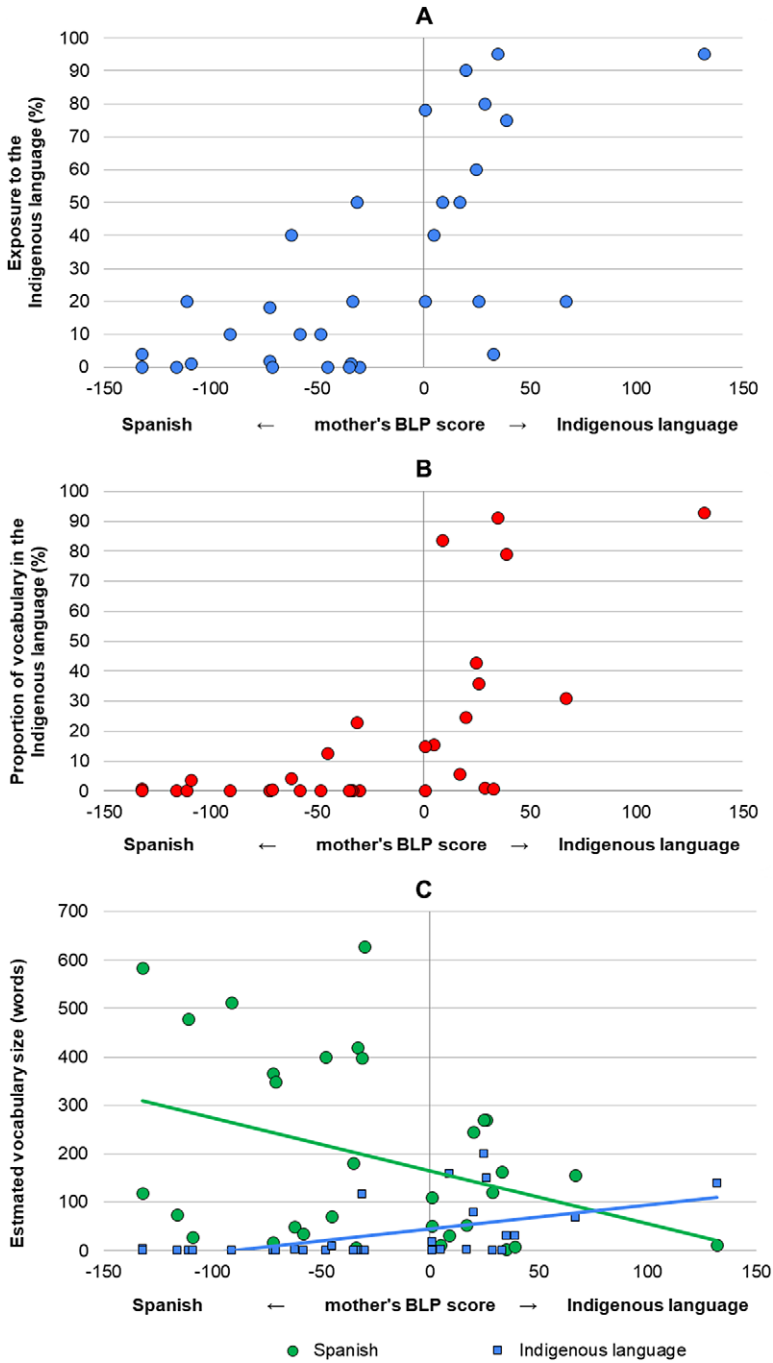

Figure 5. The effects of mothers' bilingual profile on infants' exposure to the Indigenous language (panel A), the proportion of Indigenous language vocabulary (panel B), and the infants' estimated vocabulary size in both languages (panel C).

As for the effects of age, our results showed that infants' age correlated with all vocabulary size measures except for Indigenous vocabulary size. This suggests that, while Indigenous infants' productive Spanish vocabulary will most certainly increase with age, growing up in Mexico in an Indigenous family does not necessarily guarantee that a child will acquire a productive vocabulary in the Indigenous language. This is consistent with earlier studies that show that the acquisition of a minority language in bilingual children might plateau (i.e., not further develop in time) if they are only exposed to it in the home (Mancilla-Martinez & Vagh, 2013; Miękisz et al., 2016). Instead, there are factors other than age that seem to influence infants' productive vocabulary development in the Indigenous language in this specific situation of Mexican Indigenous bilingualism.

One such factor appears to be the *amount* of exposure to an Indigenous language at home, since it is a significant predictor of the size and proportion of Indigenous vocabulary in a child's total productive vocabulary whilst controlling for age. However, the relationship between relative exposure to a language and the proportion of vocabulary in this language is not equal, since Spanish words tend to dominate infants' overall vocabulary. This is in line with previous research on minority language acquisition, where more relative exposure has been shown to be necessary in a minority language than in the majority language for the two languages to be acquired evenly (Pearson et al., 1997; Vihman et al., 2006). Our finding also corroborates the notion that, even if there is exposure, minority languages can vanish from families in as few as three generations of speakers (Fishman, 1966), and here specifically due to a language shift towards Spanish-dominant bilinguals, heritage speakers of Mexican Indigenous languages, and Spanish monolinguals with Indigenous progenitors (Canuto Castillo, 2015; Mulík et al., 2021b). These results draw parallels between Spanish acquisition within an English-dominant context (United States) and Indigenous language acquisition within a Spanish-dominant context (Mexico), in a sense that the minority status of Indigenous languages in Mexico can make it difficult for children to learn them because they are often surrounded by the majority language, Spanish, which is more prestigious and has more learning resources available to them. In the same vein, Hoff (2021) reports that bilingual children from immigrant families in Florida (with a minority language used more at home and with English as a majority, societal language) most commonly develop stronger skills in the majority language but more varied and weaker skills in the minority language. The author concludes that this is because successful language acquisition requires substantial and continued environmental support; the fact that bilingual environments do not always provide the same levels of quality and quantity of children's exposure to each language puts the minority language in disadvantage. This, indeed, seems to be the case also in the present study.

In terms of comparisons with studies on Indigenous children vocabulary development, our results are similar to those of Reese et al. (2018). The authors reported greater vocabulary sizes in English than in Māori for bilingual children from New Zealand. This resembles our results in terms of Mexican Indigenous children's bilingual vocabulary being skewed towards Spanish at the group level, which seems to be a common feature of Indigenous bilingual vocabularies in situations of Indigenous language endangerment. Additionally, Reese et al. (2018) found a negative correlation between their English and Māori vocabulary ($r = -0.49$, $p < 0.001$; Reese et al., 2018). Although in the present study we did not observe a significant correlation between the vocabulary sizes in the two languages ($r_{Spearman}(32) = -0.193$, $p = 0.291$), our results suggest a certain degree of complementarity between the Indigenous language and Spanish vocabulary. These findings are in line with the complementarity principle (Grosjean, 2013), which states

that bilinguals use different languages in different situations; it is rare for bilinguals to use all languages in all situations.

One way in which we observed the abovementioned complementarity between the two languages is via relative language dominance of the infant's mother in Spanish and the Indigenous language. Crucially, not only does the mothers' BLC score correlate with the relative exposure to the Indigenous language that the infants get, but it also correlates with their Indigenous language vocabulary size. These results suggest that measuring their primary caretaker's bilingual language profile, along with infants' age, could inform us about the infant's bilingual vocabulary development. In line with this finding of ours, Ezeizabarrena and García Fernández (2023) showed that, besides relative language exposure to Basque, the linguistic profile of parents and the language of interparental communication was also a relevant predictor of children's productive vocabulary size in Basque. However, they did not measure parents' linguistic profile on a scale as we did, but as a categorical variable instead: whether only one or both parents were speakers of Basque. As shown in our data, using a continuous bilingualism scale instead of a categorical bilingualism scale might provide a finer predictor of productive vocabulary size in bilingual children with one minority and one majority language.

Finally, in the present study, we did not observe a statistically significant effect of SES on Indigenous infants' early bilingual vocabulary development. The explanation for this result is that, in our sample, there was not sufficient SES variability captured in order for us to observe any effects of SES on early vocabulary learning (i.e., on the scale of 0–300, our samples comprised SES scores of 29–152, one family corresponding to very low [0–47], 7 to low [48–94], 16 to low-high [95–115], 5 to medium-low [116–140], and 3 to medium [141–167] social classes in Mexico). In other words, despite the data spanning five out of the total of seven social classes established in Mexico, there seems to be no effect of SES on early bilingual vocabulary size in this range (or, at least, in our sample). Nevertheless, even in the absence of such effects, our study still provides valuable data on vocabulary size and composition in bilingual Indigenous infants from low-SES contexts and its relationship to relative language exposure and mothers' bilingual language profile, with most variability being represented by the different language combinations considered in this work and various geographical location of the participants within the country.

5. Conclusions, limitations, and future studies

The present study has successfully provided novel data on early bilingual vocabulary development in Mexican Indigenous infants from low SES. In this vulnerable population, the effects of relative language exposure and mothers' language dominance on infants' vocabulary size and composition were established. Thus, the present work extends these effects, reported in literature for higher SES and non-Indigenous languages, to a common but severely understudied context of young Indigenous bilinguals of Mexico.

As mentioned above, the present study has several limitations. First of all, the instrument that was used to estimate infants' vocabulary size both in Spanish and in the Indigenous language is meant to be used for Mexican Spanish only. In the absence of a better measure of vocabulary size, we opted for using this instrument in both languages of the bilingual infants and mothers. Nevertheless, future studies would benefit from adaptations of the CDI II, specifically for the Indigenous languages and cultures of Mexico. These should account for words that may be early acquired in the Indigenous

language but not in Spanish, as well as for word frequency in each of the target languages. Such culturally and linguistically appropriate adaptations are already in place for several Indigenous languages, mainly in Africa and Oceania (see https://mb-cdi.stanford.edu/adaptations.html). Second, in the present work, we measured language exposure as relative exposure to the Indigenous language compared to the exposure to Spanish. In future studies, absolute exposure to the languages might also be explored as a factor, since infants from different cultures (be they monolingual or bilingual) might be receiving different amounts of language input as they grow. The vast cultural diversity of Mexico makes this point especially relevant for studying early vocabulary development in Indigenous bilingual children. Also, herein only the productive vocabulary was considered; however, receptive vocabulary size should also be measured in future work. This would be especially relevant for passive bilinguals, that is, bilinguals who understand both languages but speak only one of them. Finally, future studies that aim to collect CDI or other vocabulary data in Indigenous languages such as those considered in the present work should rely on close collaboration with community researchers (i.e., members of the Indigenous community that can be trained in data collection and analysis). In our experience, this aspect was crucial not only for participant recruitment and for establishing effective communication with Indigenous mothers, but also in order to understand the cultural and linguistic aspects of each visited community and language. It also makes comprehensive communicating of the outcomes of such studies to the communities more feasible.

The present work also has implications for education and Mexican Indigenous bilingualism in general. According to our results, infants' mothers' bilingual language profile seems to be a key variable that influences the exposure to the Indigenous language from home, as well as the children's vocabulary size in the Indigenous language and their relative proportion of Indigenous words in their production vocabulary. Indigenous communities could benefit from these findings to help mitigate intergenerational language loss. In face of the continuous decline of the Indigenous language use among the native populations of Mexico, a revalorisation of Indigenous bilingualism as an asset rather than a hurdle is needed.

Funding statement. S.M. would like to thank CONAHCYT for financial support (Grant No. BP-PA-2022061913162847l-2306121).

Competing interest. The authors declare no competing interests. No AI tools beyond spell-check and English grammar correction were used in the preparation of the manuscript.

References

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *The American Psychologist*, **63**(7), 602–614. https://doi.org/10.1037/0003-066X.63.7.602

Asociación Mexicana de Agencias de Inteligencia de Mercado y Opinión AC. (2022). *Nivel Socioeconómico AMAI 2022*. https://amai.org/descargas/Nota_Metodologico_NSE_2022_v5.pdf

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M., & **Gillam, R. B.** (2012). The measure matters: Language dominance profiles across measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, **15**(3), 616–629. https://doi.org/10.1017/S1366728912000090

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Birdsong, D., Gertken, L. M., & **Amengual, M.** (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL – The University of Texas at Austin.

Bleses, D., **Makransky, G.**, **Dale, P. S.**, **Højen, A.**, & **Ari, B. A.** (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, **37**(6), 1461–1476. https://doi.org/10.1017/S0142716416000060

Bosch, L., & **Ramon-Casas, M.** (2014). First translation equivalents in bilingual toddlers' expressive vocabulary: Does form similarity matter? *International Journal of Behavioral Development*, **38**(4), 317–322. https://doi.org/10.1177/0165025414532559

Byers-Heinlein, K., **Gonzalez-Barrero, A. M.**, **Schott, E.**, & **Killam, H.** (2024). Sometimes larger, sometimes smaller: Measuring vocabulary in monolingual and bilingual infants and toddlers. *First Language*, **44**(1), 74–95. https://doi.org/10.1177/01427237231204167

Canuto Castillo, F. (2015). Otomíes en la ciudad de México. La pérdida de un idioma en tres generaciones. *Lengua y migración*, **7**(1), 53–81.

CDI. (2016). *Fichas de información básica de la población indígena, 2015*. https://www.gob.mx/inpi/articulos/fichas-de-informacion-basica-de-la-poblacion-indigena-2015

CONEVAL. (2019). *La pobreza en la población indígena de México, 2008–2018*. https://www.coneval.org.mx/Medicion/MP/Paginas/Pobreza_Indigena.aspx

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)

Dale, P. S., & **Penfold, M.** (2011). *Adaptations of the MacArthur–Bates CDI into non-US English languages*. https://mb-cdi.stanford.edu/documents/AdaptationsSurvey2011.pdf

David, A., & **Wei, L.** (2008). Individual differences in the lexical development of French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, **11**(5), 598–618. https://doi.org/10.1080/13670050802149200

De Cat, C. (2021). Socioeconomic status as a proxy for input quality in bilingual children? *Applied Psycholinguistics*, **42**(2), 301–324. https://doi.org/10.1017/S014271642000079X

De Houwer, A., **Bornstein, M. H.**, & **Putnick, D. L.** (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, **35**(6), 1189–1211. https://doi.org/10.1017/S0142716412000744

DeAnda, S., **Arias-Trejo, N.**, **Poulin-Dubois, D.**, **Zesiger, P.**, & **Friend, M.** (2016). Minimal second language exposure, SES, and early word comprehension: New evidence from a direct assessment. *Bilingualism: Language and Cognition*, **19**(1), 162–180. https://doi.org/10.1017/S1366728914000820

Dowling, T., & **Whitelaw, E.** (2018). Working towards culturally and linguistically diverse speech assessments for South African children: A Xhosa case study. *Language Matters*, **49**(2), 39–61. https://doi.org/10.1080/10228195.2018.1467479

Duursma, E., **Romero-Contreras, S.**, **Szuber, A.**, **Proctor, P.**, **Snow, C.**, **August, D.**, & **Calderón, M.** (2007). The role of home literacy and language environment on bilinguals' English and Spanish vocabulary development. *Applied Psycholinguistics*, **28**(1), 171–190. https://doi.org/10.1017/S0142716406070093

Ezeizabarrena, M. J., & **García Fernández, I.** (2023). Tres indicadores de exposición a la lengua y su efecto en la adquisición del léxico de una lengua minoritaria. *Revista de Logopedia, Foniatría y Audiología*, **43**(2), 100289. https://doi.org/10.1016/j.rlfa.2022.04.003

Falk, E. B., **Hyde, L. W.**, **Mitchell, C.**, **Faul, J.**, **Gonzalez, R.**, **Heitzeg, M. M.**, **Keating, D. P.**, **Langa, K. M.**, **Martz, M. E.**, **Maslowsky, J.**, **Morrison, F. J.**, **Noll, D. C.**, **Patrick, M. E.**, **Pfeffer, F. T.**, **Reuter-Lorenz, P. A.**, **Thomason, M. E.**, **Davis-Kean, P.**, **Monk, C. S.**, & **Schulenberg, J.** (2013). What is a representative brain? Neuroscience meets population science. *Proceedings of the National Academy of Sciences*, **110**(44), 17615–17622. https://doi.org/10.1073/pnas.1310134110

Farah, M. J. (2017). The neuroscience of socioeconomic status: Correlates, causes, and consequences. *Neuron*, **96**(1), 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

Farkas, G., & **Beron, K.** (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, **33**(3), 464–497. https://doi.org/10.1016/j.ssresearch.2003.08.001

Fenson, L., **Marchman, V. A.**, **Thal, D. J.**, **Dale, P. S.**, **Reznick, J. S.**, & **Bates, E.** (2012). *MacArthur–Bates communicative development inventories, second edition* [dataset]. https://doi.org/10.1037/t11538-000

Fernald, A., **Marchman, V. A.**, & **Weisleder, A.** (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, **16**(2), 234–248. https://doi.org/10.1111/desc.12019

Fishman, J. A. (1966). *Language loyalty in the United States: The maintenance and perpetuation of non-English mother tongues by American ethnic and religious groups*. Mouton & Company.

Floccia, C., **Sambrook, T.**, **Delle Luche, C.**, **Kwok, R.**, **Goslin, J.**, **White, L.**, **Cattani, A.**, **Sullivan, E.**, **Gervain, J.**, **Abbott-Smith, K.**, **Krott, A.**, **Mills, D.**, **Rowland, C.**, & **Plunkett, K.** (2018). Vocabulary of 2-year-olds learning English and an additional language: Norms and effects of linguistic distance. I: Introduction. *Monographs of the Society for Research in Child Development*, **83**(1), 7–29.

Frank, M. C., **Braginsky, M.**, **Yurovsky, D.**, & **Marchman, V. A.** (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, **44**(3), 677–694. https://doi.org/10.1017/S0305000916000209

Gatt, D., & **O'Toole, C.** (2016). Risk and protective environmental factors for early bilingual language acquisition. *International Journal of Bilingual Education and Bilingualism*, **20**(2), 117–123. https://doi.org/10.1080/13670050.2016.1179926

Gertken, L. M., **Amengual, M.**, & **Birdsong, D.** (2014). Assessing language dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 208–225). Multilingual Matters.

Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean, P. Li, & E. Bialystok (Eds.), *The psycholinguistics of bilingualism*. John Wiley & Sons.

Hammer, C. S., **Lawrence, F. R.**, & **Miccio, A. W.** (2008). Exposure to English before and after entry into head start: Bilingual children's receptive language growth in Spanish and English. *International Journal of Bilingual Education and Bilingualism*, **11**(1), 30–56. https://doi.org/10.2167/beb376.0

Hart, B., & **Risley, T. R.** (1995). *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes Publishing Co.

Henrich, J., **Heine, S. J.**, & **Norenzayan, A.** (2010). Most people are not WEIRD. *Nature*, **466**(7302), Article 7302. https://doi.org/10.1038/466029a

Hoff, E. (2003). Causes and consequences of SES-related differences in parent-to-child speech. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 147–160). Lawrence Erlbaum

Hoff, E. (2006). Language experience and language milestones during early childhood. In K. McCartney, & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 232–251). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470757703.ch12

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, **49**(1), 4–14. https://doi.org/10.1037/a0027238

Hoff, E. (2021). Why bilingual development is not easy. In *Advances in Child Development and Behavior* (Vol. **61**, pp. 129—167). JAI. https://doi.org/10.1016/bs.acdb.2021.03.002

Hoff, E., **Core, C.**, **Place, S.**, **Rumiche, R.**, **Señor, M.**, & **Parra, M.** (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, **39**(1), 1–27. https://doi.org/10.1017/S0305000910000759

Hurtado, N., **Marchman, V. A.**, & **Fernald, A.** (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, **11**(6), F31–F39. https://doi.org/10.1111/j.1467-7687.2008.00768.x

INALI. (2008). Catálogo de las Lenguas Indígenas Nacionales. In *Variantes Lingüísticas de México con sus Autodenominaciones y Referencias Geoestadísticas*. Instituto Nacional de Lenguas Indígenas.

INEGI. (2011). *Principales Resultados del Censo de Población y Vivienda 2010.* https://www.inegi.org.mx/programas/ccpv/2010/

Jackson-Maldonado, D., **Bates, E.**, & **Thal, D.** (2003). *MacArthur Inventario del Desarrollo de Habilidades Comunicativas. Inventario II: Palabras Y Enunciados*. Brookes.

Jackson-Maldonado, D., **Thal, D.**, **Marchman, V.**, **Bates, E.**, & **Gutierrez-Clellen, V.** (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language*, **20**(3), 523–549. https://doi.org/10.1017/S0305000900008461

Junker, D. A., & **Stockman, I. J.** (2002). Expressive vocabulary of German–English bilingual toddlers. *American Journal of Speech-Language Pathology*, **11**(4), 381–394. https://doi.org/10.1044/1058-0360(2002/042)

Kelly, B. F., **Kidd, E.**, & **Wigglesworth, G.** (2015). Indigenous children's language: Acquisition, preservation and evolution of language in minority contexts. *First Language*, **35**(4–5), 279–285. https://doi.org/10.1177/0142723715618056

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, **32**(1), 69–92. https://doi.org/10.1017/S0142716410000299

Mancilla-Martinez, J., & **Vagh, S. B.** (2013). Growth in toddlers' Spanish, English, and conceptual vocabulary knowledge. *Early Childhood Research Quarterly*, **28**(3), 555–567. https://doi.org/10.1016/j.ecresq.2013.03.004

Mancilla-Martinez, J., **Pan, B. A.**, & **Vagh, S. B.** (2011). Assessing the productive vocabulary of Spanish–English bilingual toddlers from low-income families. *Applied Psycholinguistics*, **32**(2), 333–357. https://doi.org/10.1017/S0142716410000433

Marchman, V. A., **Martínez-Sussmann, C.**, & **Dale, P. S.** (2004). The language-specific nature of grammatical development: Evidence from bilingual language learners. *Developmental Science*, **7**(2), 212–224. https://doi.org/10.1111/j.1467-7687.2004.00340.x

Meir, N., & **Armon-Lotem, S.** (2017). Independent and combined effects of socioeconomic status (SES) and bilingualism on children's vocabulary and verbal short-term memory. *Frontiers in Psychology*, **8**, 1442. https://doi.org/10.3389/fpsyg.2017.01442

Miękisz, A., **Haman, E.**, **Łuniewska, M.**, **Kuś, K.**, **O'Toole, C.**, & **Katsos, N.** (2016). The impact of a first-generation immigrant environment on the heritage language: Productive vocabularies of Polish toddlers living in the UK and Ireland. *International Journal of Bilingual Education and Bilingualism*, **20**(2), 183–200. https://doi.org/10.1080/13670050.2016.1179259

Moriguchi, Y., & **Shinohara, I.** (2019). Socioeconomic disparity in prefrontal development during early childhood. *Scientific reports*, **9**(1), 2585. https://doi.org/10.1038/s41598-019-39255-6

Mulík, S., **Amengual, M.**, **Maldonado, R.**, & **Carrasco-Ortíz, H.** (2021a). Hablantes de herencia: ¿una noción aplicable para los indígenas de México? *Estudios de Lingüística Aplicada*, **73**(2), 7–37. https://doi.org/10.22201/enallt.01852647p.2021.73.970

Mulík, S., **Corona-Dzul, B.**, **Amengual, M.**, & **Carrasco-Ortíz, H.** (2021b). Perfil psicolingüístico de los bilingües otomí (hñäñho)-español, migrantes de Santiago Mexquititlán a Santiago de Querétaro, México. *Cuadernos de Lingüística de El Colegio de México*, **8**, e154. https://doi.org/10.24201/clecm.v8i0.154

O'Toole, C., & **Hickey, T. M.** (2019). Bilingual language acquisition in a minority context: using the Irish–English Communicative Development Inventory to track acquisition of an endangered language. *In Risk and protective environmental factors for early bilingual language acquisition* (pp. 30–46). Routledge.

Parra, M., **Hoff, E.**, & **Core, C.** (2011). Relations among language exposure, phonological memory, and language development in Spanish–English bilingually developing 2-year-olds. *Journal of Experimental Child Psychology*, **108**(1), 113–125. https://doi.org/10.1016/j.jecp.2010.07.011

Patterson, J. L. (2004). Comparing bilingual and monolingual toddlers' expressive vocabulary size. *Journal of Speech, Language, and Hearing Research*, **47**(5), 1213–1215. https://doi.org/10.1044/1092-4388(2004/089)

Patterson, J. L., & **Pearson, B. Z.** (2004). Bilingual lexical development: Influences, contexts, and processes. In B. A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish–English speakers* (pp. 77–104). Paul H. Brookes Publishing Co.

Pearson, B. Z. (2007). Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics*, **28**(3), 399–410. https://doi.org/10.1017/S014271640707021X

Pearson, B. Z., & **Fernández, S. C.** (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning*, **44**(4), 617–653. https://doi.org/10.1111/j.1467-1770.1994.tb00633.x

Pearson, B. Z., **Fernández, S. C.**, & **Oller, D. K.** (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, **43**(1), 93–120. https://doi.org/10.1111/j.1467-1770.1993.tb00174.x

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**(1), 41–58. https://doi.org/10.1017/S0142716400009863

Petitto, L. A., **Katerelos, M.**, **Levy, B. G.**, **Gauna, K.**, **Tétreault, K.**, & **Ferraro, V.** (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language*, **28**(2), 453–496. https://doi.org/10.1017/s0305000901004718

Place, S., & **Hoff, E.** (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2½-year-olds. *Bilingualism: Language and Cognition*, **19**(5), 1023–1041. https://doi.org/10.1017/S1366728915000322

Poulin-Dubois, D., **Bialystok, E.**, **Blaye, A.**, **Polonia, A.**, & **Yott, J.** (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, **17**(1), 57–70. https://doi.org/10.1177/1367006911431198

Pye, C. (2021). Documenting the acquisition of indigenous languages. *Journal of Child Language*, **48**(3), 454–479. https://doi.org/10.1017/S0305000920000318

Reese, E., **Keegan, P.**, **Mcnaughton, S.**, **Kingi, T. K.**, **Carr, P. A.**, **Schmidt, J.**, **Mohal, J.**, **Grant, C.**, & **Morton, S.** (2018). Te Reo Māori: Indigenous language acquisition in the context of New Zealand English. *Journal of Child Language*, **45**(2), 340–367. https://doi.org/10.1017/S0305000917000241

Samuelson, L. K. (2021). Toward a precision science of word learning: Understanding individual vocabulary pathways. *Child Development Perspectives*, **15**(2), 117–124. https://doi.org/10.1111/cdep.12408

Singh, L., **Yeung, W.-J. J.**, **Cheng, Q.**, & **Heng, E. Y.-T.** (2023). The home literacy environment mediates effects of socio-economic status on infant vocabulary development. *Developmental Science*, **26**(4), e13349. https://doi.org/10.1111/desc.13349

Snowling, M. J., **Duff, F. J.**, **Nash, H. M.**, & **Hulme, C.** (2016). Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *Journal of Child Psychology and Psychiatry*, **57**(12), 1360–1369. https://doi.org/10.1111/jcpp.12497

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, **15**(4), 426–445. https://doi.org/10.1177/1367006911403202

Thordardottir, E., **Rothenberg, A.**, **Rivard, M.-E.**, & **Naves, R.** (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*, **4**(1), 1–21. https://doi.org/10.1080/14769670500215647

Tooley, U. A., **Bassett, D. S.**, & **Mackey, A. P.** (2021). Environmental influences on the pace of brain development. *Nature Reviews Neuroscience*, **22**(6), 372–384. https://doi.org/10.1038/s41583-021-00457-5

Vihman, M. M., **Lum, J. a. G.**, **Thierry, G.**, **Nakai, S.**, & **Keren-Portnoy, T.** (2006). The onset of word form recognition in one language and in two. In P. McCardle, & E. Hoff (Eds.), *Childhood bilingualism* (pp. 30–44). Multilingual Matters. https://doi.org/10.21832/9781853598715-004

Vogt, P., **Mastin, J. D.**, & **Aussems, S.** (2015). Early vocabulary development in rural and urban Mozambique. *Child Development Research*, **2015**, e189195. https://doi.org/10.1155/2015/189195

Cite this article: Mulík, S., & Arias-Trejo, N. (2025). Bilingual Vocabulary Development in Mexican Indigenous Infants: The Effects of Language Exposure from Home and Mothers' Language Dominance. *Journal of Child Language* 1–25, https://doi.org/10.1017/S0305000924000667