results\_markdown

ASM

13/01/2023

## Methods

### Participants

Participants were recruited through convenience sampling when they visited Concordia Infant Research Laboratory to participate in a study. Families were contacted from a database of interested families in Montréal, Canada, largely via provincial birth lists, social media, and in-person recruitment, for example at libraries and community events. Our data consists of demographic and language exposure information collected between the years 2013 and 2020 via parental questionnaires and forms (described below). Caregivers filled out the questionnaires during each visit as part of the lab standard practice. At each visit, the infants also participated in several different experimental tasks, but these are not the focus of this research. All parents signed a consent form, and they were given a small thank-you gift for their participation. The current study was approved by Concordia University Human Research Ethics Board (Certification Number 10000439).

The original sample consisted of data from 552 participants. This sample included repeated measures from some families who visited the lab twice (142), three times (54), and four or more times (8). However, for our main analyses we decided to only keep one observation per family, thus We kept families who contributed data once as well as the data from the first visit of families who contributed data during multiple visits. The final data thus consisted of 439 participants aged 4.34 to 30.95 (M=13.83, SD=6.56). Of whom 211 were female. We further excluded 66 children who heard more than 10% of a third language.

While for the main analyses we excluded the second visit data of families’ who came to multiple visits, we constructed a secondary data set (n= 204) which included first and second visit data from families who contributed data during multiple visits. We used this secondary data set to explore and describe our longitudinal data. We however excluded the third and fourth visit data from both the main and the exploratory analyses because there were very few data points ().

### Instrument

Information concerning the language environment of the infant was gathered through the Language Exposure Questionnaire (LEQ; Bosch & Sebastián-Gallés, 2001; See appendix A) following the Multilingual Approach to Language Estimates (MAPLE; Byers-Heinlein et al, 2019), in which interviewers walk caregivers through a series of questions designed to help them accurately remember and realize the language environment and exposure of their infant.

The first set of questions for the LEQ with MAPLE interview aim to understand how caregivers use their languages in daily life, particularly when talking to their baby. These questions use a likert type of scale to describe the frequency with which parents use each language when talking to their child (regularly, sometimes, or never). These first set of questions helped us determine the family language strategies, for example if a caregiver uses French regularly and never uses English while the other caregiver does the opposite, we assume they are using a one-parent-one-language strategy.

The first set of questions for the LEQ with MAPLE interview aim to give a set of estimates of the amount of time babies spend hearing each of their languages throughout their lives based on a calculation made by the interviewer from caregiver responses (overall estimate), a rough estimate based on an estimation by the caregivers themselves (xxx estimate), and an overall estimate that averages the interviewer and the caregiver estimations (global estimate). The exposure estimates made with LEQ with MAPLE have high reliability (r = .77–.97) when compared to transcriptions of daylong home recordings, which are a direct measure of the infants’ linguistic environment (Orena et al., 2018). We additionally collected demographic information about our participants and their families via a short paper form.

## Results

### Data Pre-Processing

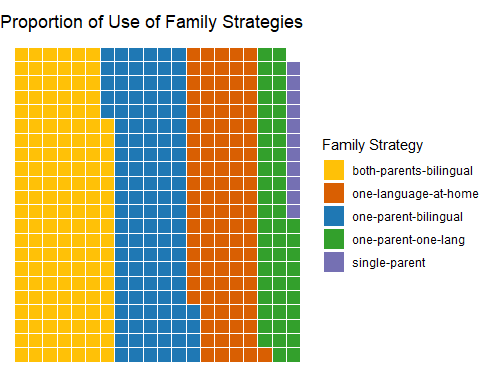
All the relevant LEQ data were manually entered from LEQ forms into excel spreadsheets. The data were entered twice by two different researchers to to catch errors, and discrepancies were resolved by a third individual who checked the original paper form. Demographic data were exported from Filemaker.

We used the responses from the first section of the LEQ questionnaire to characterize the family language strategy used by each family. For example, for a family where both parents indicated they used both their first and second language regularly we assigned the both parents bilingual strategy. Whereas for a family where each parent spoke a different first language regularly we assigned the one-parent-one-language strategy. In this manner we identified four distinct strategies: both parents bilingual, one parent bilingual, one-parent-one-language, and one language at home (see Table 1). We had originally sub-divided the one-parent-one-language strategy into a strict and flexible categories. In these sub-categories parents that always spoke their first language regularly and never spoke another language would be categorized as strict, and parents who always spoke their first language regularly but sometimes would speak another language were categorized as flexible. However, we noticed both categories were indistinct when performing statistical tests and other exploratory analyses, and thus we ultimately decided to keep them as a single group. Finally, we identified a few families with a single caregiver and we characterized them as single parent families, as we didn’t have enough cases to identify distinct strategies which single parents might use.

Table 1

### Analytic approach

We first described the family strategies (See Figure 1). The most common strategy was the both parents bilingual strategy which represented 33.9% of the sample.

Figure 1 

We performed a chi-square goodness of fit test to determine weather the difference in proportion of usage between strategies was statistically significant, where the null hypothesis is that the strategies are used in the same proportion across the families tested. The chi-squared goodness of test revealed a statistically significant difference between the distribution of proportions in our data and a distribution in which all the proportions are equal (X= 138.92, p= <.001). We were also interested in weather the mean exposure to the majority languages French and English, and the mean exposure to a minority language changed as a function of which family language strategy was used (Figure 2)

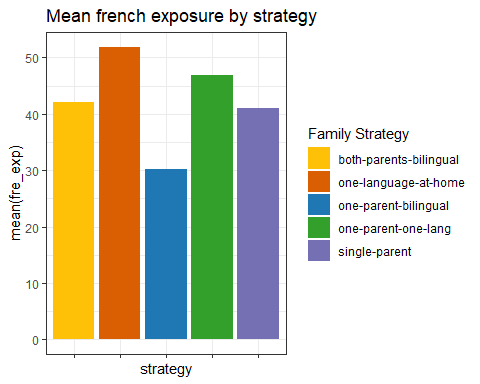
Figure 2 a) 

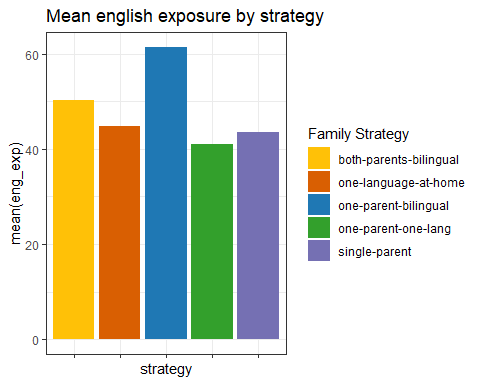
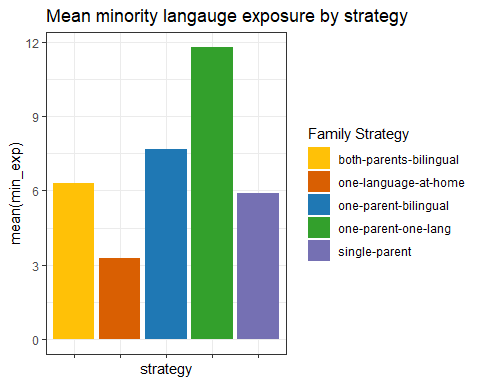
Figure 2 b) 

Figure 2 c) 

To test this statistically, we conducted a one-way analysis of variance (ANOVA test) to explore weather some family language strategies contributed more variance to the amount of exposure babies got in French, English and/or a heritage language.

In terms of French exposure, we found a statistically significant difference in how much variance the family language strategies contribute (F=7.4452, NA, p= 0.0001, NA). We then conducted a Hukey HSD post-hoc analysis to determine which observed difference between means accounted for more variance(See Table 2). The post-hoc analyses showed that the one parent bilingual strategy was statistically significantly different than the both parents bilingual, one-parent-one-language, and the one language at home strategies. The one parent bilingual strategy was associated with the lowest mean French exposure out all all the strategies (see Figure 2a).

In terms of English exposure, we found a statistically significant difference in how much variance the family language strategies contribute (F=5.5454, NA, p= 0.001, NA). We then conducted a Hukey HSD post-hoc analysis to determine which observed difference between means accounted for more variance(See Table 3). The post-hoc analyses showed again that the one parent bilingual strategy was statistically significantly different than the one-parent-one-language, and the one language at home strategies. However, in the case of English, the one parent bilingual strategy was associated with the highest mean English exposure out all all the strategies (See Figure 2b).This makes sense as most of the families in our sample were French-English bilinguals, and thus English exposure was in most cases a complement to French exposure.

Finally, in terms of exposure to a heritage language, we found an approaching statistically significance difference in how much variance the family language strategies contribute (F=2.5736, NA, p= 0.0536, NA). The fact that the the difference between strategies only approaches significance might be due to a lack of power, since the sample of families who used a heritage language was relatively small. We thus decided to also conduct a Hukey HSD post-hoc analysis. The post-hoc analysis showed that the one language at home strategy was statistically significantly different than the one-parent-one-language strategy. The one language at home strategy was associated with the lowest mean heritage language exposure, while the one-parent-one-language strategy was associated with the highest mean heritage language exposure (see Figure 2c).

Table 2

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| term | group1 | group2 | null.value | estimate | conf.low | conf.high | p.adj | p.adj.signif |
| strategy | BothParentsBilingual | OneLanguageAtHome | 0 | 9.771 | -1.913 | 21.4538 | 0.1370 | ns |
| strategy | BothParentsBilingual | OneParentBilingual | 0 | -11.999 | -23.334 | -0.6636 | 0.0332 | \* |
| strategy | BothParentsBilingual | OneParentOneLang | 0 | 4.820 | -9.926 | 19.5663 | 0.8340 | ns |
| strategy | OneLanguageAtHome | OneParentBilingual | 0 | -21.769 | -34.053 | -9.4860 | 0.0000 | \*\*\*\* |
| strategy | OneLanguageAtHome | OneParentOneLang | 0 | -4.950 | -20.437 | 10.5365 | 0.8430 | ns |
| strategy | OneParentBilingual | OneParentOneLang | 0 | 16.819 | 1.593 | 32.0450 | 0.0237 | \* |

##Table 3

kable(tukey\_eng)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| term | group1 | group2 | null.value | estimate | conf.low | conf.high | p.adj | p.adj.signif |
| strategy | BothParentsBilingual | OneLanguageAtHome | 0 | -5.513 | -17.5358 | 6.509 | 0.6380 | ns |
| strategy | BothParentsBilingual | OneParentBilingual | 0 | 11.177 | -0.4869 | 22.841 | 0.0659 | ns |
| strategy | BothParentsBilingual | OneParentOneLang | 0 | -9.252 | -24.4254 | 5.922 | 0.3950 | ns |
| strategy | OneLanguageAtHome | OneParentBilingual | 0 | 16.691 | 4.0509 | 29.331 | 0.0040 | \*\* |
| strategy | OneLanguageAtHome | OneParentOneLang | 0 | -3.738 | -19.6742 | 12.198 | 0.9300 | ns |
| strategy | OneParentBilingual | OneParentOneLang | 0 | -20.429 | -36.0964 | -4.761 | 0.0046 | \*\* |

##Table 4

Finally, we wanted to explore weather families switched in their use of family language strategies across time.To explore this, we used a subset of of our data that included families that had provided LEQ data two separate visits (i.e. longitudinal data). As mentioned in the method section we had a few families that had contributed data to three, four or more data points, however these families were so few we decided not to include data beyond the second point. We wxplored the longitudinal data (See Figure 3) to determine thew stability of family strategy use across time.

Figure 3 