MAPLE: A Multilingual Approach to Parent Language Estimates

Krista Byers-Heinlein, Esther Schott, Ana Maria Gonzalez-Barrero, Melanie Brouillard, Daphnée Dubé, Amel Jardak, Alexandra Laoun-Rubenstein, Meghan Mastroberardino, Elizabeth Morin-Lessard, Sadaf Pour Iliaei, Nicholas Salama-Siroishka and Maria Paula Tamayo

Concordia University

Bilingual infants vary in when, how, and how often they hear each of their languages. Variables such as the particular languages of exposure, the community context, the onset of exposure, the amount of exposure, and socioeconomic status are crucial for describing any bilingual infant sample. Parent report is an effective approach for gathering data about infants’ language experience. However, its quality is highly dependent on how information is elicited. This paper introduces a Multilingual Approach to Parent Language Estimates (MAPLE). MAPLE promotes best practices for using structured interviews to reliably elicit information from parents on bilingual infants’ language background, with an emphasis on the challenging task of quantifying infants’ relative exposure to each language. We discuss sensitive issues that must be navigated in this process, including diversity in family characteristics and cultural values. Finally, we identify six systematic effects that can impact parent report, and strategies for minimizing their influence.

Introduction

Many infants around the world grow up in bilingual, multilingual, and bi/multidialectal settings. Describing and quantifying these language environments is both important and complex. While there are a variety of approaches, including daily diaries (De Houwer, 2011; Place & Hoff, 2011) and home recordings (De Houwer & Bornstein, 2016; Marchman, Martínez, Hurtado, Grüter & Fernald, 2016), parent report continues to be a widely-used method for assessing young children’s language background (DeAnda, Bosch, Poulin-Dubois, Zesiger & Friend, 2016; Liu & Kager, 2016; Paradis, Emmerzael & Duncan, 2010; Unsworth, 2016).

Parent report has many advantages: it is relatively fast and easy to implement. Moreover, it is possible to assess language exposure both currently and over the infant’s entire life, which is not usually feasible in daily diary or home recording studies. This is important given that infants’ language exposure changes with their daily routines over time. At the same time, parents’ memories, effects of social desirability, and challenges in accurately translating parents’ qualitative knowledge of their infants’ lives to a quantitative measurement can affect the reliability and validity of parent report.

Capitalizing on our lab’s extensive experience with such issues, this paper introduces a Multilingual Approach for Parent Language Estimates (MAPLE). MAPLE is a set of best practices for sensitively and accurately gathering information about infants’ language background. While we will discuss the specific questionnaire our lab uses as an example, we note that MAPLE should be implemented with consideration for the specific bilingual context. Moreover, while we have most often used MAPLE with bilingual infants, it is generalizable to infants learning any number of languages.

Key descriptors of an infant’s language environment

Bilingual infants’ language environments vary along a myriad of dimensions, and it is likely that many of these affect language acquisition. Here, we put forward four key descriptors that we believe are essential in characterizing the language background of bilingual samples, regardless of the research question: the languages of exposure, the community context, the onset of exposure, and the amount of exposure. These variables have been consistently reported in studies examining bilingual infants’ language and cognitive development (e.g., Bosch & Sebastián-Galles, 1997; Byers-Heinlein, Fennell & Werker, 2013; Poulin-Dubois, Blaye, Coutya & Bialystok, 2011; Sundara, Polka & Molnar, 2008). The next sections will discuss each of these key descriptors in turn.
Socioeconomic status (SES) is a fifth key descriptor that we recommend be routinely measured and reported in studies of bilingual infants (see Figure 1 for an overview of all descriptors). Children from high SES families tend to have better language outcomes than children from low SES families (Dollaghan, Campbell, Paradise, Feldman, Janosky, Pitcairn & Kurs-Lasky, 1999; Fernald, Marchman & Weisleder, 2013; Hoff, 2003). Importantly, SES can vary systematically between monolinguals and bilinguals in the same community (Morton & Harper, 2009), and between different populations of bilinguals, making SES information vital for interpreting group differences. We refer readers to discussions elsewhere of approaches for measuring SES (e.g., Bornstein, Hahn, Suwalsky & Haynes, 2014; Shavers, 2007).

There are other input variables that we have not identified as essential to report about every bilingual sample. These include: who is providing the input (e.g., parents, siblings, daycare; Gathercole, 2014; Unsworth, 2016), the number of speakers of each language (Gollan, Starr & Ferreira, 2014), the language proficiency of different caregivers (e.g., their fluency and accentedness; Place & Hoff, 2011), and whether the language environment has a one-person-one-language structure (Döpke, 1992; although see Byers-Heinlein, 2014; Goodz, 1989 for evidence that this approach is rarely strictly followed). There is emerging evidence that these variables contribute to bilingual children’s language development (Byers-Heinlein, 2013; De Houwer, 2007; Gollan et al., 2014; Place & Hoff, 2016). However, their role and how they should be measured are much less well understood than the key descriptors we have identified. We consider these additional variables to be of high interest for studies specifically focused on understanding the role of different types of input. However, we argue that, for now, they are not essential to report about all samples of bilingual infants.

Languages of exposure

While it might seem trivial, it is important to describe the specific languages to which a sample is exposed. For example, a study could report that infants were exposed to both Spanish and Catalan (a homogeneous sample), or that infants were exposed to English and to a second language that varied, while listing those languages (a heterogeneous sample). It is often informative to specify the dialect or variety of the language (e.g., Quebec or European French, the specific variety of Chinese, etc.). This information is essential for understanding the nature of infants’ experiences (e.g., knowing whether their languages are closely related), and how this might affect study outcomes.

Community context

Even for children learning the same language pair, some grow up in bilingual communities where many adults speak both languages, while others grow up in monolingual communities where one of their languages receives limited support outside the home. Given extensive evidence that the status of a language in the community can affect language learning outcomes (i.e., whether it is a majority or a minority language; De Houwer, 2007; Huguet, 2007; Smithson, Paradis & Nicoladis, 2014), describing bilingual infants’ community context is vital for understanding children’s wider linguistic and educational milieu, as well as interpreting research findings. Unlike other variables that are specific to the child, community context will be similar for infants living in the same geographic area. Information about official languages and languages used in the community is usually available on government websites.

Onset of exposure

Bilingual infants vary in when they start hearing their second language. Simultaneous bilinguals are exposed to both languages during their first days of life (De Houwer, 1990; Werker & Byers-Heinlein, 2008), with both languages considered first languages (Meisel, 1989; 2004). Other infants begin acquiring a second language some months after having been exposed to a single language from birth, which we would consider to be early sequential bilinguals. Note that this definition of simultaneous versus sequential differs from a common arbitrary cutoff used in the adult literature, in which age 3 is often set as a dividing line between simultaneous and sequential bilinguals (e.g., Klein, Mok, Chen & Watkins, 2014). The onset of exposure is relevant because language acquisition can be seen as a system of cascading critical periods (Werker & Hensch, 2015), where input (or lack of input) from particular languages can affect language development at later stages. Studies should report – in as much detail as possible – the onset of exposure to each language, and whether this

---

**Fig. 1.** Key variables that should be measured and reported when describing samples of bilingual infants. Other descriptors may be informative for specific studies depending on the research question.
was constant (e.g., all simultaneous bilinguals) or varied (e.g., infants were exposed sometime before 12 months).

**Amount of exposure**

While some infants have relatively equal exposure to their different languages, others have more exposure to one of them (often called the dominant language). The amount of exposure to each language can influence infants’ performance on language tasks (e.g., Byers-Heinlein, Morin-Lessard & Lew-Williams, 2017; Hurtado, Grüter, Marchman & Fernald, 2014; Ramon-Casas, Swingley, Sebastián-Gallés & Bosch, 2009). The amount of exposure can refer either to an infant’s current or to their cumulative (lifetime) exposure to each language, which will likely differ if the infant’s caregiving situation has changed. Research suggests that since cumulative and current exposure can make different contributions to language development (Unsworth, 2012), measuring both is recommended. In addition, researchers vary considerably in the minimum amount of regular exposure required to consider an infant bilingual (Byers-Heinlein, 2014). Accordingly, it is important to have a clear and consistent operational definition of bilingualism that is determined prior to data collection and analysis. Researchers should report both the mean and range of exposure to each language.

**Eliciting information on key descriptors from parent report**

Parents are good informants about the key descriptors of infants’ language background, as they are familiar with the people who spend time with their infant and the languages they speak. It is typically fairly straightforward to determine which languages an infant is exposed to, the community context, and when the exposure began. However, it is more complicated to accurately assess the amount of exposure to each language.

An obvious approach is to simply ask parents directly, for example “What percent of the time does your child hear English versus French?” While this approach is quick and efficient, it has significant limitations. First, parents often find it difficult to think about percentages in relation to their child’s language exposure. Second, when such a question comes out of the blue, parents may not thoroughly consider all of the different factors that contribute to their infant’s language exposure, and may be strongly influenced by reporting biases (Schwarz, 2007). Parents may also think differently from researchers about what counts as an important source of input: a point we will return to later.

Under MAPLE, we advocate using a structured interview to overcome some of these issues. Our structured interview lasts about 15 minutes (depending on the child’s age and history) and guides parents through a series of questions about their infant’s language background and everyday life to quantify the infants’ language exposure. MAPLE also emphasizes the importance of the interviewer establishing rapport with parents, explaining what “counts” as exposure, answering questions that may arise when parents think about their child’s language experience, and adapting questions to the infants’ particular situation and background. This structured nature improves the reliability of measurement across different individuals administering the questionnaire. While quantifications of language exposure between single questions and structured interviews are correlated, we have noted that structured interviews can reduce the effects of reporting biases (Byers-Heinlein, 2017). We will discuss effects that bias parent report on language environments in more detail in a later section.

**Using a structured interview to assess infant language experience**

This section describes our process for conducting structured interviews to assess infant language background, an approach pioneered by Bosch and Sebastián-Gallés (2001). One of us (Byers-Heinlein) has been conducting these types of interviews for more than a decade, and together we have conducted thousands of such interviews with parents from a variety of backgrounds, while at the same time refining what would become MAPLE over the years. We recommend this procedure for infants from a few months old until around 3 years, considering that later in development children’s language use and proficiency becomes increasingly important and measurable alongside their language input (e.g., Bedore, Peña, Summers, Boerger, Resendiz, Greene, Bohman & Gillam, 2012; Bohman, Bedore, Pena, Mendez-Perez & Gillam, 2010). MAPLE can be adapted to measure older children’s input (i.e., asking about experience year-by-year rather than month-by-month), but additional measures of children’s proficiency and use will be needed.

Our structured interview is designed to provide comprehensive information about an infant’s language environment: the languages of exposure, the onset of exposure, and the amount of exposure. Community context usually depends on where the research is taking place, and thus usually does not need to be elicited from parents. Our aim is to quantify both current and cumulative exposure to each language the infant hears. We use a pen-and-paper questionnaire with a structured script, downloadable from the Open Science Framework at osf.io/bvxfz (see also DeAnda et al., 2016, for an Excel macro that can automate some calculations). It is ideal to interview parents in their preferred language, although this might not always be possible. We describe the entire interview in detail in the sections to follow, and have provided videos of sample interviews with completed questionnaires on our Open Science Framework page.

**Background questions**

The first part of the interview assesses the infant’s languages and sources of input. The interviewer begins by asking about the language(s) spoken by the infant’s caregivers, both in their daily lives as well as specifically to the child. Caregivers could include parents, other relatives, nannies, babysitters, or family friends who spend an hour per week or more with the infant on a regular basis. Age of acquisition, language variety, and degree of non-native accent are also elicited for each main caregiver. While we do elicit these detailed variables about the language background of caregivers as they might sometimes be of interest, once again we do not consider them essential for every study to report. We also gather information concerning language use at daycare and other activities outside the home, as well as any long family trip in which the infant’s language exposure might have changed. Importantly, the information obtained during this first part of the interview allows the interviewer to gather more accurate and detailed answers concerning amount of exposure in the next section.

**Day-in-the-life**

The second part of the interview walks parents through a typical day in the infant’s life, which enables them to estimate the number of hours the infant is exposed to each language per day and
explain where this exposure comes from. The interviewer begins by determining the number of hours the infant is awake per day. For simplicity, we usually assume that this is constant across the infant’s life, given limitations in parents’ ability to recall detailed information about infants’ past sleeping patterns. Parents are then prompted to think about who the child interacts with during those waking hours and in which languages, usually going through the day chronologically, to determine how many hours the infant typically hears each language. For practical reasons, we set the minimum time unit of interest to half an hour per day.

In order to assess both cumulative and current exposure, as well as onset of exposure, this walk-through is done separately for each period of the infant’s life in which their daily routine – and likely their language exposure – differed. Thus, the interviewer starts from the infant’s first month of life, asks questions to determine the daily exposure to different languages at birth, and then asks at what point there might have been a change in the infant’s exposure. This process is repeated in increments of one month or more up to the infant’s current age. The interviewer determines any differences in language exposure between weekdays and weekends as well. This method ensures that changes in language exposure across the child’s life are taken into account, including: parents’ return to work after leave, starting daycare, changes that occur in response to the child’s preferred language, long vacations, time spent with different individuals, etc.

A crucial step in this part of the interview is explaining to parents what type of input is taken into account as exposure. A large body of literature suggest that speech directed to the child, rather than television or overheard speech, has the greatest effect on early language processing and vocabulary (Hoff, 2006; Kostyrka-Allchorne, Cooper & Simpson, 2017; Kuhl, Tsao & Liu, 2003; Weisleder & Fernald, 2013). We thus ask parents to report only language spoken directly to the child and exclude all other sources such as television and overheard speech. However, we note that future research might point to new ways of considering this indirect input.

Global measure

The last step of the interview elicits a single global measure of the child’s language exposure from parents. At this point, the parent has been thinking deeply about the infant’s language environment and has a clear idea of what counts as input. The interviewer prompts parents to estimate the percentage of words heard by the child in each language throughout his or her whole life. Here, parents may be able to tap into a more global sense of their infant’s exposure, accounting for days that are not “typical”, which might not be captured in the day-in-the-life approach. This global measure is used when calculating the final overall estimate of language exposure.

Final estimate of language exposure

Using the number of hours per day of exposure during each life period obtained from the day-in-the-life interview, the interviewer computes an estimate of the percentage of time the infant has heard each of the languages over their lifespan – the cumulative exposure – as well as the infant’s current exposure. Importantly, for the cumulative exposure, this calculation is weighted by the number of months in each period (e.g., a 10 month period will have twice the weight as a 5 month period). Last, to obtain the final estimate of language exposure, we average the global measure provided by the parents with the cumulative exposure obtained from our calculations. This combines the computational approach of the day-in-the-life questions, with the “Gestalt” approach of the global measure, in order to calculate a more reliable final estimate. Note that the day-in-the-life and the global measures are expected to be similar, and a large discrepancy between them should prompt the interviewer to clarify the source of the inconsistency. Preliminary work shows that this final estimate of language exposure using MAPLE has high reliability (r = .77–.97) when compared with transcriptions of daylong home recordings (Orena, Byers-Heinlein & Polka, 2019).

Interviewer sensitivity

Conducting a structured interview to assess an infant’s language environment naturally brings an interviewer in contact with considerable information about the child’s family structure, culture, and upbringing. Sensitivity and openness to different family situations and practices is essential for establishing trust between an interviewer and the family in order to gather accurate information about a child’s language environment.

Family characteristics

Families differ considerably in terms of who regularly cares for the child. For instance, children may be raised by two parents of different sexes or of the same sex, by a single parent, by parents living in different households, or in households that include extended family members. Some infants have non-family caregivers (e.g., daycare, nanny) while others do not. It is therefore important to be mindful of different family compositions and caregiving arrangements, and speak in general terms if details are not known (e.g., “Does your child have another caregiver?” is preferred over “What about your child’s father?”). Since we are interested in assessing language exposure from the beginning of the infant’s life, we need to gather any information involving changes in family structure or caregiving, especially if these resulted in changes in language exposure. Some of these changes can be difficult topics to bring up, such as separation or divorce, unemployment, or the death of a caregiver. We have found that, when probing gently, parents are usually very open to discussing what consequences these events might have had on their child’s language exposure.

Cultural differences

Interviewing parents who speak different languages typically involves interacting with individuals from various cultures and ethnic groups, and interviewers should be sensitive to cultural issues that arise when interviewing parents. Certain cultural practices may lead to changes in an infant’s life. For example, in some Chinese families, an infant born in North America may spend some months living in China with his/her grandparents without the infant’s parents. Interviewers should enquire in general terms on any changes that occurred in the child’s environment, rather than making assumptions about what is considered a typical upbringing.

Sensitivity in talking about language(s)

Cultures may also differ in terms of the labels they use for their languages. For example, some parents report speaking
“Chinese” to their child. However, “Chinese” is a general term that could refer to different varieties such as Mandarin, Cantonese, or Hokkien, amongst others. Just like a child learning French and Spanish is bilingual, children learning Mandarin and Cantonese are considered bilingual as well. The use of informal terms such as “Chinese” or “Creole” should cue an interviewer to ask additional questions to find out more about which language variety an infant is hearing. It is also important for interviewers to use terminology that demonstrates a sensitivity to the family’s culture. Interviewers should mirror the parents’ own terminology when possible, and use specific terms for languages when necessary.

Effects that may influence parent report

Just like self-report, parent report can be influenced by systematic reporting biases and effects (Schwarz, 2007). In the next sections, we identify six effects that we encounter frequently when interviewing parents about their infant’s language background, and discuss ways to mitigate their influence.

Good Parent Effect

“I want my child to be bilingual, so that means he/she is hearing English half of the time and Spanish half of the time.” All parents want to provide the best possible platform for their children’s growth and development, and they might have particular ideas about the language environments that are most beneficial. For instance, in the Canadian context of our lab, bilingualism is seen as highly desirable, and we have observed that parents may specifically highlight their child’s bilingual language exposure. This makes it important to distinguish between parents’ intentions and beliefs about what is best for their child, and the child’s actual language experience. Interviewers can mitigate the Good Parent Effect by using neutral language that holds all languages, dialects, and language environments in equal regard. Walking parents through an actual day in the life of their infant during the structured interview also helps to gauge the child’s language input accurately.

Equal Caregiving Effect

“I only speak French and my partner only speaks English, so our child is exposed to 50% French and 50% English.” In the case where parents follow a one-parent-one-language approach, some parents may assume that this leads to equal input in both languages. The reality is that, most of the time, one caregiver spends more time with the infant, often the mother in the early months. In these cases, the maternal language will likely be the dominant language. It is then important to avoid asking questions that imply that the input from one caregiver is more important than from another, and rather ask about how many hours each caregiver interacts with the child.

Dora Effect

“My child hears a lot of Spanish from watching Dora the Explorer every day.” Parents may think that infants learn languages from watching television shows or videos. A popular example at the time of writing is Dora the Explorer, a television show that includes simple Spanish words and phrases. As previously discussed, we exclude exposure to television from our assessment of language exposure, which may be contrary to some parents’ beliefs about how language is learned in infancy. However, if a parent mentions that Dora – or any other television show/character – has provided a substantial proportion of an infant’s experience in a language, the interviewer can take note of the television exposure, but exclude it in further quantification of the child’s language exposure.

Grocery Store Effect

“My child hears a lot of French from everybody who says ‘bonjour!’ at the grocery store.” The Grocery Store Effect stems from a belief that very brief interactions in the community provide children with a significant amount of language exposure. However, for most children, these experiences will only add up to a few minutes a day. To overcome the Grocery Store Effect, the interviewer can specify the smallest unit of language exposure that is of interest (e.g., half an hour), and ask parents whether the child hears more than that on a daily basis. If the exposure is minimal, this helps parents put the actual input into perspective.

Wishful Grandma Effect

“We have lunch with Grandma every week, and she speaks both English and Italian, so my child knows Italian too.” Many parents want their children to learn their family’s heritage language, and in some families, grandparents or other extended family members might be the primary source of input in that language. This could lead parents to overestimate the importance of the grandparents’ language in their child’s life, especially if they do not live in the home with the child. In that event, the interviewer can probe more specifically about how many hours a week the child spends with extended family members to gauge the child’s exposure to that language. It can also be helpful to remind parents that the day-in-the-life estimate only includes language spoken directly to the child, as in some cases extended family members speak a language largely amongst themselves but only rarely directly to the child.

Daycare Black Hole Effect

“It’s supposed to be a bilingual daycare, but I don’t know if this is really the case.” If an infant attends daycare where the educators are bilingual, it is often difficult for parents to estimate their child’s language exposure at daycare. Parents may have some idea from observing the educators’ interactions with their child during dropoff and pickup. If there are several educators at daycare who speak different languages with the children, it may also be useful to ask the parents whether their child spends more time with a particular educator. In principle, researchers could send questionnaires to the child’s daycare educator, but answers might, as discussed above, be subject to cognitive biases, and trying to ensure that the questionnaires are returned to the researcher is a considerable logistical effort. Although it may not be perfect, parents often remain our best source of information, and the interviewer can ask parents to give their best estimate of the language exposure provided at daycare.

Conclusion

Understanding young children’s language environments is critical for contextualizing their language development, and is particularly relevant for infants growing up bilingual or multilingual.
Parent report is one of the primary ways researchers can assess infants’ language background. MAPLE encompasses a wide variety of specific practices, and here we have highlighted our lab’s method as one way that a structured interview can elicit key information from parents and caregivers. Optimal strategies for implementing MAPLE may differ somewhat from lab to lab, depending on linguistic and cultural factors unique to the population of interest. Nevertheless, all researchers will benefit from carefully considering what needs to be measured in their population, how to elicit complete and accurate information from parents, the ways in which the interviewer can be sensitive to a particular family and culture, and systematic effects that can influence parent report.

Author ORCIDs. Krista Byers-Heinlein, 0000-0002-7040-2510; Esther Schott, 0000-0002-0114-4773; Ana Maria Gonzalez-Barrero, 0000-0002-2120-6329; Melanie Brouillard, 0000-0002-7129-3221; Elizabeth Morin-Lessard, 0000-0001-7218-1611.

Acknowledgements. This work was supported by a Natural Sciences and Engineering Research Council of Canada grant to KBH, by undergraduate, graduate, and post-doctoral awards from the Fonds de Recherche du Québec - Société et Culture to AMGB, ES, EML, AJ, and SPE; from the Centre for Research on Brain, Language and Music to EML; and from Concordia University to ES, SPE, and DD. KBH holds the Concordia University Research Chair in Bilingualism.

References


