Is mommy talking to daddy or to me? Exploring parental estimates of child language exposure using the Multilingual Infant Language Questionnaire

Liquan Liu & René Kager

To cite this article: Liquan Liu & René Kager (2017) Is mommy talking to daddy or to me? Exploring parental estimates of child language exposure using the Multilingual Infant Language Questionnaire, International Journal of Multilingualism, 14:4, 366-377, DOI: 10.1080/14790718.2016.1216120

To link to this article: https://doi.org/10.1080/14790718.2016.1216120
Is mommy talking to daddy or to me? Exploring parental estimates of child language exposure using the Multilingual Infant Language Questionnaire

Liquan Liu\textsuperscript{a,b} and René Kager\textsuperscript{b}

\textsuperscript{a}School of Social Sciences and Psychology, Western Sydney University, Sydney, Australia; \textsuperscript{b}Utrecht Institute of Linguistics-OTS, Utrecht University, Utrecht, The Netherlands

**ABSTRACT**

Language input is a key factor in bi-/multilingual research. It roots in the definition of bi-/multilingualism and influences infant cognitive development since and even before birth. The methods used to assess language exposure among bi-/multilingual infants vary across studies. This paper discusses the parental report patterns of the amount/degree of exposure to their children and provides an algorithm-based Multilingual Infant Language Questionnaire (MILQ) targeting the amount of hours and degree of exposure an infant is exposed to each language. In the MILQ, parental feedback between general language input (languages spoken in the environment an infant resides in) and direct language exposure (languages spoken directly to an infant) are differentiated. Comparing the results drawn from general and direct perspectives, parental estimates of their children’s exposure match the general but not direct language input condition. Implications of these results are discussed.

**Introduction**

Input plays a significant role in child development. Previous infant studies have been incongruent in the ways of approaching and discussing this factor. From a conceptual perspective, infants face various types of linguistic input in their ambient environment. Apart from infant-directed speech, an infant may receive language input by interacting with adults, siblings, and even peers. Interpersonal interaction is essential for language acquisition, especially in the first year of life (Kuhl, Tsao, & Liu, 2003). Nevertheless, studies targeting the extent to which certain input types may influence language acquisition are rare. At a representational level, various methods have been adopted across studies on multilingualism\textsuperscript{1} to assess input, yet few studies have offered in-depth discussions on the source and nature of the collected input data.

In the following paragraphs, a brief overview of the measurements of multilingual input information across studies is provided, followed by a discussion between general and direct input. The discussion on the type of input leads to our main question how input...
is estimated by families with multilingual children. Thereafter, the notion of amount/degree of exposure (AoE/DoE) is highlighted, with data collection and comparisons on parental estimates versus the DoE of multilingual infants on general and direct language input calculated by the Multilingual Infant Language Questionnaire (MILQ) designed by the author.

**Measuring the language exposure of a multilingual child**

It is difficult to measure the absolute amount of language input an infant is exposed to in her daily life, and this is even more so for infants from multilingual backgrounds. Some studies use recording devices, such as the Language Environment Analysis system (LENA™, Gilkerson & Richards, 2008, 2009), to assess children’s linguistic environment. With its small size and lightweight, the digital language processor of LENA can be placed in a pocket next to a child and record up to 16 hours of continuous speech data, illustrating the language exposure of the child.

LENA is an excellent tool for the understanding of the quantitative and qualitative linguistic input of a child. Nevertheless, its current form does not suit all types of language studies. First, although it is possible to record the input from surroundings periodically (e.g. two weeks), one must make estimates if she aims at applying the input measured from LENA to a longer time frame (e.g. annually) during which a child’s systematic schedule changes may occur. It is thus hard to compare the accuracy of LENA’s outcomes with that of other input measurements, all of which by far require a certain degree of estimation. Second, to meet the requirement of the conversation analysis by LENA, data from children’s speech output are necessary. In other words, it is difficult to accurately verify child-directed speech (language input directly spoken to a child) when the child is not responding (e.g. inattentive) or cannot respond to a conversation (e.g. physical or vocabulary constraints in the first year after birth). Third, LENA is primarily designed for English learning environment. Its validity on data collection of languages other than English is being examined in recent studies (e.g. Canault, Le Normand, Foudil, Loundon, & Thai-Van, 2015). Fourth, and perhaps most importantly, LENA’s current acoustic analytical software does not allow it to easily adapt to the diverse multilingual environments, recognise and separate individual language experience of a child.

Apart from online information collection through speech recording devices, parental reports on their children’s linguistic ability, such as the MacArthur-Bates Communicative/Developmental Inventory (Fenson et al., 1993), have widely been used in infant research. Compared to those from children’s pre-school teachers, language reports from children’s parents show higher accuracy (Vagh, Pan, & Mancilla-Martinez, 2009). Parents of a multilingual child can also estimate each language their child encounters in her complex environment, although how good their estimations are remains unclear. A multilingual environment is commonly a complex one. Children may be exposed to multiple languages at the same time, and parents may mix multiple languages when speaking to each other or to their child since these adults are likely to speak multiple languages themselves. In the latter case, language mixing may occur at the sentence, word, or even phoneme level. In a survey investigating the language mixing situation among 181 bilingual families in Canada through parental reports, only 4% of the families stick to the ‘one-parent-one-language’ strategy 100% of the time and merely 14% families adhere to this strategy 90% of the
time (Byers-Heinlein, 2013). In addition, a multilingual child’s choice in speech production has been shown to be related to input statistics and social interaction (Ghimenton, 2013; Ghimenton, Chevrot, & Billiez, 2013). The diversity of language use in different environments imposes unneglectable challenges: a child’s exposure to each language in the home environment may differ from that at daycare or other social environments. Although studies on multilingual vocabulary development illustrate that parents are able to report how many words their child understands or speaks across ages (De Houwer, Bornstein, & Putnick, 2014; Pearson, Fernández, & Oller, 1993), to what extent the complex multilingual environment may create difficulty for parents to estimate the language exposure for their infants remains unclear. It is urgent to investigate how good parents are at reporting the linguistic input of their multilingual child.

The importance of the direct and indirect speech

The complex multilingual environment introduces a crucial issue that is often neglected: parental reports normally do not distinguish infant general language exposure (languages spoken in the environment an infant resides in) from a speech direct setting (languages spoken directly to her). A distinction should be made between infant general and direct language exposure. The formula below addresses the relationship between the two types of exposure:

\[
\text{General language experience} = \text{direct exposure} + \text{indirect exposure}.
\]

In this paper, direct input is defined as the input provided through direct interaction with an infant, whereas indirect input refers to the input that is spoken in the surroundings but not directly to the infant. One’s general language experience is characterised as the sum of the two types of input.

This difference between direct and indirect input is not trivial and sheds light on mechanisms of early language acquisition. On the one hand, children less than two years of age can learn from speech not directly addressed to them (Oshima-Takane, 1988; Oshima-Takane, Goodz, & Derevensky, 1996). On the other hand, it has been argued that the direct but not the indirect input is essential for learning and development among young infants (Pearson, Fernández, Lewedge, & Oller, 1997; Shneidman, Todd, & Woodward, 2014). Infant-directed speech, for instance, facilitates language acquisition in various aspects (Kuhl, 2004). Yet, it remains unclear whether direct language input plays a more significant role than indirect input in infant language acquisition. Although direct input has been shown to be crucial for early word learning even in communities where a sizable or majority proportion of children’s early language input comes from overheard speech (e.g. Shneidman, Arroyo, Levine, & Goldin-Meadow, 2013; Yucatec Mayan children, Shneidman & Goldin-Meadow, 2012), such effect is by no means automatic and innate (Shneidman & Woodward, 2015).

Given the differences in influence between distinct types of input in infancy, it is, therefore, important to consider these distinctions when collecting the input data. This issue is more addressed in recording methods like LENA than in other data collection methods. LENA analyses and distinguishes ‘Adult Words Report’, an estimation of the number of adult words spoken to and near the child, and ‘Conversational Turns Report’, the number of adult–child conversational interactions. The former report matches the
above-mentioned general language condition, and the latter is close to a direct input situation. The issue, however, is that such adult–child conversational interactions cannot fully represent the actual amount of languages directly addressed to an infant since she may not illustrate a response feedback before 12 months of age or even after. The issues as to which type of exposure parents tend to report and how accurate they are in reporting the information remain unanswered in previous parental reports and questionnaires collecting the input information.

When discussing direct versus indirect speech, factors like birth order is worth mentioning. Whereas firstborn infants may experience more one-to-one interaction with a caretaker, later-born infants may be exposed to increased indirect speech from older siblings (Shin, 2002). The potential influence of birth order/sibling should be noted by input measurement tools in multilingual research.

**The impact of DoE on language development**

Despite the issue of the type of exposure among multilingual infant studies, parental reports measure a child’s DoE (more often than AoE) to each of the languages. AoE/DoE and language dominance are essential for multilingual infant studies that measure input as an independent factor. AoE/DoE usually refers to the (estimated) absolute amount/relative degree of each language a multilingual infant is exposed to, whereas the definition of language dominance varies. It may refer to the language commonly used in the home environment, the language that possesses social dominance, or the language with a higher AoE/DoE among all languages in the ambient environment. The current paper focuses on the measurement of DoE.

The DoE is not only used to evaluate or select multilingual population but also appeared occasionally in monolingual studies. Since multilingualism has become the new norm in the 21st century, certain limitation of DoE to a second language is often needed to constrain the monolingual population. A strong relationship exists between multilingual infants’ AoE/DoE to a language and their linguistic competence in that language (Hoff, 2006; Martínez, Rodríguez, Marchman, Hurtado, & Fernald, 2013; Pallier, Colomé, & Sebastián-Gallés, 2001; Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009; Thordardottir, 2011). For example, the DoE to a language influences the phonemic and phonotactic development of a multilingual child (Ramon-Casas et al., 2009; Sebastián-Gallés & Bosch, 2002), as well as her neural responses to language from infancy to adulthood (Garcia-Sierra et al., 2011; Pallier et al., 2001).

The DoE standards and measurements vary across studies. Some studies adopt a DoE in the non-dominant language as low as 9% (Marchman, Fernald, & Hurtado, 2010), whereas others as high as 35% (Bosch & Sebastián-Gallés, 2003). There is no standardised DoE percentile value to define a ‘balanced’ or ‘unbalanced’ multilingual, although less than 20% of DoE of a language has been argued to lead to a passive use of that language (Pearson et al., 1993). As for the measurement, some early studies mention that participants have significant exposure to both languages on a regular basis without measuring their DoE (Pearson, Fernández, & Oller, 1995). Others use minimum hours of exposure per week, varying from 8 to 20 hours, for the non-dominant language (Junker & Stockman, 2002; Patterson, 1998).

Several questionnaires have been designed and used in previous multilingual research; yet, not all questionnaires focus on the type of input and DoE in a multilingual
environment. Bosch and Sebastián-Gallés (1997) use a language exposure questionnaire, asking questions about the languages spoken by all caregivers who are present since the birth of the infant, as well as parents’ overall estimation of their infant’s language exposure. Byers-Heinlein (2013) designs a language mixing scale questionnaire specifically targeting the effect of language mixing within a bilingual family, and studies the influence of parents’ language mixing on their infants’ language development. Unsworth (2013) uses accumulated measurement to estimate the previous language exposure of Dutch–English bilingual children. Previous questionnaires often use categorically based questions, leading to a rough estimation of DoE. A rough or categorical estimation of DoE appears to be less accurate when comparing the effect of language exposure and dominance within the multilingual population, making it difficult to study the correlation effect between DoE of a language and language proficiency of that language.

This paper investigates parents’ DoE estimation of their infants with two aims. First, we illustrate parental estimation patterns and robustness through the MILQ. Second, we demonstrate the difference between general (direct and indirect) versus direct input and parental estimation patterns of their multilingual infants, calling for awareness in input types and multilingual selection criteria. The research questions are: What are parents’ estimation patterns when estimating the DoE of their multilingual infants? Is the pattern closer to the general or direct input their infants hear? How well do parents estimate?

Experiment
Instrument
The MILQ, designed and adopted in our previous studies (Liu, 2014; Liu & Kager, 2013, 2015a, 2016; in press), can apply to a wide range of multilingual settings from infants aged 0 to 3 years. It is used in this study to capture both general and direct exposure. Specifically, the MILQ calculates multilingual infants’ hours of exposure to each language in the main locations and/or situations of their daily lives: babysitting, daycare, home, and social environment (Figure 1). Some locations are optional depending on the life routine of individual infants. Taking home environment as an example, the language information of each person living with the infant is recorded. Parents fill in the average waking hours infants stay in these locations, the percentage of time each language is spoken during stay, and the percentage of time when they interact directly with their infants. Since parents of a multilingual child often come from different countries, their language exposure when travelling abroad is included. The general and direct DoE is generated by comparing the sum of the hours of each language an infant is exposed to or directly spoken to in all situations/locations. The MILQ provides parents with a comprehensive view of their infants’ language backgrounds. When a family has more than one child growing up in a similar environment, the older child’s speech output, functioning as an additional source of language exposure to her/his younger sibling, is captured by the MILQ.

Participants
Families with typically developing 5–15-month-old bilingual infants participated in the study. Socio-economic status affects early language development (Feldman et al., 2000;
Hoff, 2003). Due to the high quality of the social welfare system in the Netherlands, the socio-economic status differs little across participating families. In terms of the educational background, at least one parent from each family obtains a degree equivalent to university or higher level. All bilingual infants were exposed to Dutch as one of their native languages, and the other language varied across participants. In the first test, parents from 48 families filled in the MILQ and estimated their infants’ language exposure and both general and direct input exposure hour in each environment. In the second test, data of 109 families were included in the general input part of the questions of the MILQ.

**Procedure**

Parents filled in the questionnaire together with the experimenter in the babylab of Utrecht Institute of Linguistics OTS, Utrecht University. In the beginning of the MILQ, parents were required to estimate the DoE to each language of their infants based on their intuition without any initial bias or knowledge of the general, direct or indirect input (parental estimation). The question was: ‘What is the percentage of each language the baby hears?’ Upon the completion of the MILQ, infants’ DoE information was generated based on an algorithm calculating the hours of exposure to each language considering all situations/locations (general DoE, direct DoE).

**Figure 1.** The main location/situation components of the MILQ.
Results

We compared the outcomes of the Dutch exposure since it is the only language shared across participants. Paired samples \( t \)-tests were conducted in between parents’ initial Dutch DoE estimation of their children, general DoE, and direct DoE calculated by the MILQ. The mean (SD) of Dutch DoE is listed in Table 1. A significant difference was found between parental estimation and direct DoE, \( t(1, 47) = 2.260, \ p = .028 \) (2-tailed), but not between parental estimation and general DoE, or between general and direct DoE (\( p > .05 \)). That is, parental estimation tends to be closer to the language input from the general environment than that to which infants are directly spoken. Figure 1 demonstrates the scatterplot of individual differences between DoE and parental estimation (Figure 2).

Among the 48 families, 11 have siblings in the household. Due to the limited sample size, we do not find evidence of the birth order having an effect on parental estimation patterns. Interestingly, among 10 out of 11 families (91%) with siblings, the dominant language heard by the infant matched the dominant language spoken by the sibling.

Parental estimation was closer to general DoE and differed significantly from direct DoE. We enlarged the sample size and focused on the difference between parental estimation and general DoE to Dutch. Paired samples \( t \)-test with 109 infants reported no significant difference between parental estimation and general input calculated by the MILQ (\( p > .05 \)), replicating our initial finding. The mean and SD of Dutch DoE are listed in Table 2.

Table 1. The mean (SD) of Dutch DoE in each estimation.

<table>
<thead>
<tr>
<th>DoE type</th>
<th>Mean Dutch DoE (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parental estimation</td>
<td>54.44% (21.30%)</td>
</tr>
<tr>
<td>General DoE</td>
<td>51.28% (21.53%)</td>
</tr>
<tr>
<td>Direct DoE</td>
<td>49.55% (21.30%)</td>
</tr>
</tbody>
</table>

Figure 2. The scatterplot of individual differences of Dutch DoE (general DoE versus parental estimation and direct DoE versus parental estimation).
Discussion

When dealing with multilingual development, parental reports concerning their children’s language exposure are an issue of importance. Our results show that parental estimation of their infants’ DoE is closer to infants’ general language exposure, the input from direct interaction and indirect environment, rather than direct interaction only.

These results provide implications relevant to our research questions. Regarding the question on parents’ estimation patterns, the MILQ data show that parents’ estimation of their infants’ DoE differs significantly from the direct but not general DoE. This indicates that parents are aware of the overall language exposure of their multilingual children in the ambient environment. Furthermore, parents’ estimation of their infants’ DoE is fairly close to the results generated from the MILQ, indicating their relatively accurate intuition of their children’s language exposure. In addition, the current results suggest that the MILQ successfully captures the language exposure of a multilingual child by measuring infants’ hours of exposure to each language in the main locations and/or situations.

The current results have important implications not only for studies using parental report but also for those exploring the effect of input. Input frequency and input social embedding play crucial roles in language acquisition (Liu & Kager, 2014). They facilitate infant learning process (Saffran, Aslin, & Newport, 1996) and influence their perceptual patterns (Jusczyk & Luce, 1994). Nevertheless, insufficient attention has been paid to the disentanglement, distinction and comparison between direct and indirect input, as well as their respective effects. The MILQ provides a viable measurement and a valuable instrument for research conducted in this field.

The current finding sheds light on previous debates whether direct but not indirect input from the ambient environment plays a significant role in early development (Oshima-Takane et al., 1996; Shneidman et al., 2014). It has been argued that in order to acquire a language, the type of input an infant needs to be exposed to must be direct rather than indirect (Pearson et al., 1997). Nevertheless, evidence supporting this argument is scarce. It has been shown that a hearing child with two deaf parents does not learn to speak or sign without direct exposure of speech or sign to them (Griffith, 1985; Sachs & Johnson, 1976). Additionally, American infants aged 9 months alter their perception of Mandarin when given systematic pre-exposure (45 minutes a week for 12 weeks) to the language via interpersonal interaction, whereas pre-exposure via audiovisual or audio-only recordings of Mandarin fails to facilitate these infants’ perception of Chinese (Kuhl et al., 2003).

Although direct input may indeed weigh more than indirect input in infant language development, indirect input also contributes to language acquisition. Children less than two years of age can learn pronominal references from indirect input exposure (Oshima-Takane, 1988). The facilitation effect of indirect exposure may be stronger in a multilingual in comparison to a monolingual environment since previous research has

| Table 2. The mean (SD) of Dutch DoE in each estimation with enlarged sample size. |
|-------------------------------|-------------------------------|
| DoE type                      | Mean Dutch DoE (SD)           |
| Parental estimation           | 54.24% (20.29%)               |
| General DoE                   | 53.44% (21.28%)               |
shown that multilingual infants are highly sensitive to social cues and social contexts (Mattock, Polka, Rvachew, & Krehm, 2010; Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012). These effects, often considered as ‘advantages’ stemming from a multilingual environment, have been reported across linguistic, cognitive, and social domains. The degree of influence from each type of input on language acquisition, as well as the potential maturational constraints that may alter such influence at different time windows along the language developmental trajectory, remains unclear. Additionally, in most multilingual studies, no clear distinction between the DoE of general and direct input has been discussed. Research targeting the influence of direct language exposure should specify the type of input collected from the multilingual environment.

Finally, we would like to list a number of important issues not discussed in the current study for future research. First, given that direct and indirect exposure may play different roles in infant language development, their respective and interactive functions under different social settings should be investigated. Second, with respect to input types, although the current study highlights the language exposure in young infants, it is foreseeable that the massive invasion of modern technology may cause increases in direct and indirect exposure which may subsequently affect early acquisition. The impact of technology on infant development may become a critical field to be studied. Third, the influence of AoE/DoE on language development is of great scientific and social importance. Follow-up studies can link individual linguistic exposure with performances from infancy (Liu & Kager, 2011, 2015b) to adulthood (Chen, Liu, & Kager, 2015, 2016; Liu, Chen & Kager, in press). Fourth, the MILQ captures detailed information of siblings and other family members living with the infant. The factor of birth order/sibling on language acquisition is not yet well understood. Future studies should investigate whether such effect will lead to differences in language development among multilingual families. Last but not least, studies are being conducted testing the validity of the MILQ by comparing the questionnaire outcomes with other existing tools measuring infant language exposure (e.g. Language diary, De Houwer & Bornstein, 2003; Place & Hoff, 2015).

Notes

1. The notion of multilingualism is used to refer to both bilingual and multilingual scenarios throughout the paper unless marked.
2. For example, all families receive substantial subsidy for childcare services from the government in the Netherlands.

Acknowledgements

We would like to thank the anonymous reviewers of International Journal of Multilingualism for their valuable comments. We sincerely thank Brendan Devers and Emma Everaert for their review. We dearly thank Sharon Unsworth for her instruction and feedback in the beginning stage of the questionnaire. We are grateful for all multilingual families in the Netherlands that volunteered to participate in this research.

Disclosure statement

No potential conflict of interest was reported by the authors.
References


