Supplementary material

Paper title: A review of questionnaires quantifying bilingual experience in children: Do they document the same constructs?

Authors: Draško Kašćelan, Philippe Prévost, Ludovica Serratrice, Laurie Tuller, Sharon Unsworth, Cécile De Cat

Section A. A list of collected questionnaires

Note: The underlined questionnaires were included in our review, while the rest were excluded for reasons outlined in the paper. Furthermore, the questionnaires marked with an asterisk were the original 13 which we included in our Google Form survey inquiring about the questionnaires used by researchers of bilingualism.

1. *Alberta Language and Development Questionnaire (ALDeQ, Paradis et al., 2010)
2. *Bilingual Language Experience Calculator (BiLEC, Unsworth, 2013)
3. *COST Action IS0804 Questionnaire for Parents of Bilingual Children (PaBiQ, Tuller, 2015)
4. *Language Background Questionnaire (De Cat, 2020)
5. *Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007)
6. *Language Exposure Assessment Tool (LEAT, DeAnda et al., 2016)
7. *Language Exposure Questionnaire (Cattani et al., 2014)
8. *Language History Questionnaire (Li et al., 2006)
10. *Teacher Questionnaire (Gutierrez-Clellen & Kreiter, 2003)
12. *The Bilingual Language Profile (Birdsong et al., 2012)
13. *Virtual Linguistic Lab Child Multilingualism Questionnaire (Blume & Lust, 2017; Yang et al., 2007)
14. Anamnese Meertaligheid (based on Blumenthal & Julien, 2000, Q-BEx Team translation from Dutch)
15. Background Questionnaire (adapted by Wilson, 2017)
16. Bilingual Dominance Scale (Dunn & Fox Tree, 2009)
17. Bilingual English-Spanish Assessment (BESA): Bilingual Input and Output Survey (BIOS) - Parent (Peña et al., 2018)
18. Bilingual English-Spanish Assessment (BESA): Bilingual Input and Output Survey (BIOS) - School (Peña et al., 2018)
19. Bilingual English-Spanish Assessment (BESA): Inventory to Assess Language Knowledge (ITALK) - At Home (Peña et al., 2018)
20. Bilingual English-Spanish Assessment (BESA): Inventory to Assess Language Knowledge (ITALK) - School (Peña et al., 2018)
21. Bilingual Patient's Profile - Children (Scharff Rethfeldt, 2012a)
22. Bilingual Questionnaire (Weber-Fox & Neville, 1996)
23. Bilingual Switching Questionnaire (Rodriguez-Fornells et al., 2012)
24. Bilingualism and Emotions Web Questionnaire (Pavlenko, 2005)
25. Bilingualism Questionnaire (Silverberg & Samuel, 2004)
27. Caregiver Language Questionnaire (de Diego-Lázaro, 2019)
28. Child Questionnaire BALED project (Prentza et al., 2017)
29. Input et Expérience dans le Développement Bilingue (INEXDEB) - Parent Questionnaire (Cohen, 2015a)
30. Input et Expérience dans le Développement Bilingue (INEXDEB) - Parent Questionnaire, Short Follow-up Version (Cohen, 2015b)
31. Interview questions for stakeholders in a Turkish school (Bahar Octur, 2009)
32. Language and Social Background Questionnaire (Anderson et al., 2018)
33. Language Background Questionnaire (Flège & MacKay, 2004)
34. Language Background Questionnaire (LBQ, McKendry & Murphy, 2011)
35. Language Background Questionnaire (Marchman & Martínez-Sussmann, 2002)
36. Language background questionnaire for heritage speakers of Spanish (Torres, 2012)
37. Language background questionnaire for Turkish/Kurdish speakers - Parent Questionnaire (Marinis, 2012)
38. Language background questionnaire for Turkish/Kurdish speakers - Child Questionnaire (Marinis, 2012)
39. Language Exposure Questionnaire (Read et al., 2020)
40. Language history questionnaire (LHQ 2.0) - A new dynamic web-based research tool (Li et al., 2014)
41. Language History Questionnaire (LHQ3) - An enhanced tool for assessing multilingual experience (Li et al., 2019)
42. Language History Questionnaire (Liu et al., 1992)
43. Language history questionnaire (Loebell & Bock, 2003)
44. Language history questionnaire (Tokowicz et al., 2004)
45. Language Input Diary (LID, De Houwer and Bornstein, 2003)
46. Language Mixing Questionnaire (Byers-Heinlein, 2013)
47. Warwick Language Project Questionnaire: An adapted questionnaire based on the LEAP-Q and The Language and Social Background Questionnaire (Rodenhurst, 2020)
48. Language Use Questionnaire - updated (Vasanta et al., 2010)
49. Language Use Questionnaire (Vasanta et al., 2010)
50. Leeds-Bradford Language Exposure Questionnaire for Adults (LeBLEQ-A, Gunning & Klepousniotou, n.d.)
52. Leeds-Bradford Language Exposure Questionnaire for Teachers (LeBLEQ-T, Gunning & Klepousniotou, n.d.)
53. Macro-sociological survey questionnaire (De Houwer, 2003)
54. MAPLE - A Multilingual Approach to Parent Language Estimates (Byers-Heinlein et al., 2019)
55. Mehrsprachigkeit in Kita und Schule, Elternbefragung (Lindauer, in prep.)
56. Multilingual und Interkulturell orientierte Anamnese - MIA - Kinder English (Scharff Rethfeldt, 2012b)
57. My baby and language (De Houwer & Bornstein, 2016)
58. PaBiQ adaptation 1 (Antonijevic-Elliott, n.d.)
59. PaBiQ adaptation 2 (Antonijevic-Elliott et al., 2020)
60. Parent Questionnaire (adapted by Arredondo, 2017)
61. Parental Questionnaire BALED project (Prentza et al., 2017)
62. Parental Questionnaire (adapted by Özturk, n.d.)
63. Parental report of the child's speech or language problems and history of speech and language problems in the family (Restrepo, 1998)
64. Participants' questionnaire (Deuchar et al., 2018)
65. Protocol for obtaining information about a BFLA context (De Houwer, 2009)
66. Protocol for obtaining information about a BFLA context (De Houwer, 2009) - a duplicate
67. Questionnaire for adult speakers and child learners of English and Russian (Lyutykh, 2012)
68. Questionnaire for Parents of Bilingual Children: Infant and Toddler Version (PaBiQ-IT, Gatt et al., 2015)
69. Questionnaire for teachers about the child's language at home and at school (Restrepo, 1998)
70. Bilingual background questionnaire for Spanish/English speakers (Montrul, 2012)
71. Questionnaire in English for heritage speakers of Spanish (Carreira, n.d.)
72. Russian Language Proficiency Test: Questionnaire (for Preschool Children) (Gagarina et al., 2010)
73. Russian Language Proficiency Test: Questionnaire (for School Children) (Gagarina et al., 2010)
74. Survey of demographic and self-identification information for heritage learners of Mexican descent (Gignoux, 2009)
75. The preschool language context and activity record (De Houwer, 2002)
76. PEGEBOS-1 (De Houwer, 2016a)
77. PEGEBOS-2 (De Houwer, 2016b)
78. PEGEBOS-3 (De Houwer, 2017)
79. Web questionnaire on bilingualism and emotion (Dewaele & Pavlenko, 2001/03)
80. Why are you learning German? For students aged 8-12 (Stracke, 2011)
81. Why are you learning German? For students aged 13-17 (Stracke, 2011)
Section B. A list of overarching constructs and components identified across reviewed questionnaires

Note: Overarching constructs are capitalised and marked in **bold**, while their components are underlined and in *italics*.

**DEMOGRAPHICS (CHILD):**

1. *Child’s name*
2. *ID*
3. *Sex*
4. *Today’s date*
5. *Date of birth*
6. *Age* – chronological age
7. *Birth order* – of the target child in relation to their siblings
8. *Age/date of arrival* – to the current country of residence / societal language (SL) country
9. *Place of birth*
10. *Current residence (location and/or length)*
11. *Previous residence (location and/or length)* - In one case, the operationalisation asked about previous residence and travel longer than one week.
12. *Nationality*
13. *Ethnicity*

**(PRE)SCHOOL INFORMATION:**

1. *School*
2. *Day care/Babysitter/Preschool* – indicate if the child goes/went to one
3. *Teacher* – name
4. *SLT* – name
5. *Grade/Class*

**DEMOGRAPHICS (SIBLINGS):**

1. *Name*
2. **Number** - Note that in some questionnaires, this can be inferred from other questions, such as name of siblings, age of siblings, etc.

3. **Sex**

4. **Date of birth**

5. **Age**

6. **Birth order**

7. **Place of birth**

8. **Age of exposure** – to languages that they speak

9. **Languages** - languages that the siblings speak and/or understand. The languages spoken/understood by siblings can sometimes be inferred from the information on their ‘Age of exposure’ (same table) or for instance from ‘Interlocutor skills/quality – Proficiency’ (Table S3) or from ‘Interlocutor skills/quality – Speaking and/or understanding’ (Table S3).

**DEMOGRAPHICS (PARENTS/CAREGIVERS/INFORMANT):**

1. **Name**

2. **Relationship to child**

3. **Sex**

4. **Contact**

5. **Age/Date of birth**

6. **Place of birth**

7. **Country of origin**

8. **Current residence**

9. **Previous residence**

10. **Length of residence in the HL country**

11. **Length of residence in the SL country**

12. **Intention to stay in the SL country**

13. **Nationality**

14. **Ethnicity**

15. **Race**

16. **Languages** - variables such as languages that they speak and/or understand, age of acquisition/first exposure, native language, etc. In one case, a questionnaire also inquires about whether formal education in each language existed.
DEMOGRAPHICS (OTHER PEOPLE/GENERIC):
1. Sex
2. Age/Date of birth
3. Place of birth
4. Country of origin
5. Current residence
6. Length of residence in the SL country

INTERVIEWER’S/ADMINISTRATOR’S DATA:
1. Name
2. Contact
3. Who is completing the questionnaire? - The interviewer was one of the options; hence, the question was classified here.

TYPE OF BILINGUALISM (CHILD):
1. List of languages – this can be any of the following: names of languages, home languages (HLs), languages other than HL and SL, languages spoken and understood, languages in order of acquisition, languages other than SL, L1, L2, L3, additional language
2. Stronger language
3. Type of exposure – this can include: parental strategies regarding which language(s) they use with the child (e.g. one parent, one language), providers/contexts of first exposure, place where the child got into contact, started learning, or started getting exposed to a specific language
4. Date/Age of exposure/acquisition - One operationalisation included a question about the age when the child became fluent in each language.

EXPOSURE (note that when a questionnaire does not explicitly distinguish between exposure and use, the questions were classified in this section):
1. Number of interlocutors – who live with the child, who spend time with the child, frequent contacts, who look after the child, who interact with child, who speak each
language to the child, number of HL children in school, most with child during the
day, who help raise the child, from whom the child hears a specific language. In
several operalisations, the questionnaires inquire about the number of interlocutors in
each or in a specific language, which could therefore also be classified under
‘Interlocutor/ Context/ Activities (language used)’ (see further below in the same
table).

2. Changes in the number of interlocutors

3. Nativeness of interlocutors – Whenever a questionnaire asked about the native
language of more interlocutors than only parents/caregivers, those operalisations
would be classified here. Otherwise, documentation of native languages of
parents/caregivers was classified in Table S1, Demographics: parents/ caregivers/
informants, Languages (or see in ‘Demographics: parents/caregivers/informant >
point 16’ above).

4. Dialect of interlocutors/exposure

5. Early exposure – normally up to a certain point in life (e.g. before the age of 4, before
preschool/school, up to the age of 8, etc.). This can include any of the following:
frequency of exposure, exposure to languages in relation to contexts (places,
activities) or interlocutors, time when the exposure began, interlocutors that the child
interacted most with during this period.

6. Changes in exposure

7. Age/Period of exposure (in relation to interlocutor/context) – age and periods since or
during which the child was exposed to a certain language and by a specific
interlocutor (e.g. parents, siblings, etc.) or in a certain place/context (e.g. home,
school, etc.). One operalisation contains sub-questions about days per week in a
daycare and hours per week in out-of-school care. Even though this is a time-unit-
based measure, it was classified here as the questions were asked in relation to the age
of exposure. Another operalisation also contains a time-unit-based component
(number of day parts per week). However, as the question inquires about exposure
since a specific period in life, it was classified here.

8. Hours per day – a time-unit-based measure of exposure to languages; this exposure
can be in relation to interlocutors, contexts, activities or a combination of these. Note
that in many cases, the questionnaires inquire about the hours per each day of the
week, which also enables us to measure how many hours per week the target child is
exposed to a specific language, interlocutor, context, or doing an activity.
9. **Hours per week** – a time-unit-based measure of exposure to languages; this exposure can be in relation to interlocutors, contexts, activities or a combination of these. Operalisations classified here were the ones which could not help us infer the hours per day of exposure, but would rather inquire about the whole week (e.g. average hours per week spent with the main caregiver and the languages used).

10. **Weeks per year** – spent in a certain environment (e.g. school)

11. **Sleeping/Waking hours**

12. **Frequency of exposure (relative)** – exposure/use of languages with interlocutors, in certain contexts or during the activities; inquired about in relative terms (e.g. adverbs, percentages). This could also include overheard speech inquired about in relative terms (e.g. when child’s interlocutors speak among themselves). In one questionnaire, two operalisations include options about language mixing. In several operalisations, whenever the scale was ‘HL, both, SL’ or ‘mostly HL, both, mostly SL’, we classified this under Interlocutor/Context/Activities (language used). However, when the scale was ‘mostly HL, both equally, mostly SL’, or ‘mostly HL, both languages to the same extent, mostly SL’, this would be classified under relative exposure.

13. **Interlocutor/Context/Activities (language used)** – language(s) used with specific interlocutors, in certain contexts, during activities or a combination of these; not inquired about in relative terms. This also includes potentially overheard languages (e.g. when child’s interlocutors speak among themselves). Two operalisations include an option about language mixing. Furthermore, two operalisations inquired about target child’s contact with children from the same ethnic group and with monolingual SL children.

14. **Language mixing** – any questions in relation to language mixing in the child’s exposure

15. **Interlocutor other** – any other questions about child’s interlocutors (e.g. demographics about their friends, are friends’ parents from the target child’s HL/SL country)

**USE** (questions that explicitly ask about the target child’s language use):

1. **Early use** – up to a certain point in child’s life

2. **Changes in use**

3. **Frequency of use (relative)** – use of languages with interlocutors, in certain contexts or during the activities; inquired about in relative terms (e.g. adverbs, percentages). In
several operalisations, whenever the scale was ‘HL, both, SL’ or ‘mostly HL, both, mostly SL’, this would be classified under Interlocutor/Context/Activities (language used). However, when the scale was ‘mostly HL, both equally, mostly SL’, or ‘mostly HL, both languages to the same extent, mostly SL’, this would be classified under relative use.

4. *Hours per day* – a time-unit-based measure of use of languages; this exposure can be in relation to interlocutors, contexts, activities or a combination of these. Note that sometimes the questionnaires inquire about the hours per each day of the week, which also enables us to measure how many hours per week the target child uses a specific language.

5. *Interlocutor/Context/Activities (language used)* – language(s) used by the child towards specific interlocutors or in certain contexts or activities; not inquired about in relative terms. One operalisation includes an option about language mixing. Moreover, one operalisation includes some basic demographics about the interlocutor.

6. *Most spoken (to)* – language or interlocutor that the child speaks the most (in)to

7. *Language mixing* – language mixing in child’s use

**EARLY SKILLS HL (CHILD) (up to a certain point in life):**

1. *Speaking*
2. *Understanding*

**EARLY SKILLS SL (CHILD) (up to a certain point in life):**

1. *Speaking*
2. *Understanding*

**CURRENT SKILLS IN THE SL (CHILD):**

1. *SL speaking/production*
2. *SL listening/comprehension*
3. *SL reading*
4. *SL writing*
5. *SL overall proficiency*

**CURRENT SKILLS IN THE HL (CHILD):**
1. **HL speaking/production**
2. **HL listening/comprehension**
3. **HL reading**
4. **HL writing**
5. **HL reading and writing** – not asked about separately
6. **HL overall proficiency**
7. **Reasons for dissatisfaction** – with the level of skills

**OTHER LANGUAGE-RELATED SKILLS/CHARACTERISTICS (CHILD):**

1. **Talkativeness**
2. **Relevant language** – using relevant language in adequate situations; included questions about switching as well
3. **Questions** – appropriate response to questions

**CAREGIVER 1’S HL SKILLS/QUALITY:**

1. **Listening/Comprehension**
2. **Speaking/Production**
3. **Overall proficiency**
4. **Reading**
5. **Writing**
6. **Errors**
7. **Accent**

**CAREGIVER 2’S HL SKILLS/QUALITY:**

1. **Listening/Comprehension**
2. **Speaking/Production**
3. **Overall proficiency**
4. **Reading**
5. **Writing**
6. **Errors**
7. **Accent**

**ADDITIONAL CAREGIVER’S HL SKILLS/QUALITY:**

1. **Listening/Comprehension**
2. Speaking/Production
3. Overall proficiency
4. Reading
5. Writing

CAREGIVER 1’S SL SKILLS/QUALITY:
1. Speaking/Production
2. Listening/Comprehension
3. Overall proficiency
4. Reading
5. Writing
6. Grammar
7. Errors
8. Accent
9. SL courses – information about SL classes/courses

CAREGIVER 2’S SL SKILLS/QUALITY:
1. Speaking/Production
2. Listening/Comprehension
3. Overall proficiency
4. Reading
5. Writing
6. Errors
7. Accent
8. SL courses – information about SL classes/courses

ADDITIONAL CAREGIVER’S SL SKILLS/QUALITY:
1. Listening/Comprehension
2. Speaking/Production
3. Overall proficiency
4. Reading
5. Writing
INTERLOCUTOR SKILLS/QUALITY (when it includes all interlocutors in the house, or any interlocutors apart from parents):

1. **Proficiency**
2. **Speaking and/or understanding**
3. **Errors**
4. **Accent**

OTHER LANGUAGE SKILLS/CHARACTERISTICS (MOSTLY PARENTS/CAREGIVERS):

1. **Talkativeness**
2. **Errors (language not specified)**
3. **Age of first acquisition** - This is the only component/operationalisation under this construct not about parents/caregivers. Rather, it inquired about 2 other adults and in-house childcare.
4. **Stronger language (parent)**

CONTEXT QUALITY (this set of questions asks explicitly about the quality):

1. **Outside home**
2. **Holiday**
3. **Other sources**

ACTIVITIES:

1. **Reading in HL and/or SL** – mostly frequency (e.g. how many times per week/year) but also other related questions, apart from reading skills, which is entered above
2. **Literacy and other in HL and/or SL** – ‘other’ here refers to a wide set of activities (e.g. TV/movies/cinema, numeracy, playing games, storytelling, etc.); this section also includes the availability of printed material in specific languages as well as the frequency of these activities (e.g. how many times per week/year)
3. **Preferred activities** – of the child
4. **Ease of learning new things (how quickly?)**
5. **Activity patterns** – how many activities can the child do at a time
6. **Extra-curricular activities** – mostly questions about their frequency (e.g. how many times per week). In one operationalisation, a questionnaire inquires about the hours per week spent on several activities. One of the activities includes a foreign language,
while another one English. As these were listed with other activities, the operationalisation was classified here rather than under exposure (hours per week). In one operationalisation, one of the options includes an initiative about child’s reading on their own. This was classified here rather than under reading above, as it was inquired about in conjunction with seven other activities.

7. **Courses/Classes** – mostly about classes, courses, instruction in HL, SL or both

8. **Travel/Holidays** – mostly frequency of travelling and sometimes length of stay.

9. **Other activities** – when the set of activities is so broad; mostly frequency of doing them

**Socioeconomic status:**

1. **Parent/Caregiver/Informant education** - One operationalisation includes a question about the education of all household members, while other two ask about siblings and 'other' in addition to parents.

2. **Parent/Caregiver/Informant occupation** - One operationalisation inquires about sibling occupation and another one about other adults.

3. **Relatives’ education and profession** - broader family (e.g. relatives in the home country or relatives in general)

4. **Free/Reduced meals**

5. **Property**

6. **Income**

7. **Holidays** - only when the questionnaire used travel and holidays to operationalise the socioeconomic status

8. **Status** – in the community or on the national level

**Developmental issues/concerns:**

1. **Hearing problems**

2. **Birth/Pregnancy complications**

3. **Developmental delay**

4. **Language problems** - any language related issues, delays, struggles, worries, concerns, frustration, problems, difficulties, or diagnosis related to language impairment/delay

5. **Academic/Learning problems**

6. **Behavioural/Social problems**
7. **Physical problems**

8. **Health problems** - One operalisation inquires about hearing, cognitive functioning, and/or language use. As it is a mixture of several components reported above, this operalisation is classified here.

9. **Other** – mostly about the diagnoses or special needs that are not necessarily language specific (e.g. autism, ADHD, need of an assistant). Not listed under any other component above.

**FAMILY HISTORY:**

1. *(Mainly) Language difficulties*

**EARLY MILESTONES:**

1. **Beginning to walk**
2. **Sitting**
3. **Babbling and first words**
4. **First short sentences** - One operalisation inquires about child's first questions.
5. **Literacy**
6. **Comparison to same age children**

**OTHER BEHAVIOUR/CHARACTERISTICS (CHILD):**

1. **Shyness**
2. **Handedness**

**ATTITUDES TOWARDS LANGUAGE / IMPORTANCE OF LANGUAGES:**

1. **Parent/Caregiver encouraging child to learn**
2. **Beliefs on bi/multilingualism** – by the child or the parents
3. **Other beliefs about language/learning** – by the child or the parents
4. **Child’s goals and expectations** - about their own educational, language or other achievements
5. **Parental goals and expectations** - about child’s educational, language or other achievements. One of the operalisations contains a list of statements, some of which could be classified under 'Parent/Caregiver encouraging child to learn' or 'Other beliefs about language/learning'.

6. *Preferred language (child)* - One operationalization inquires about the culture that the child feels closer to.

7. *Preferred language (parents)* - Two operationalizations are rather about parents’ importance (for themselves) of speaking and understanding a specific language.

**ATTRITION:**

1. *Language loss (child)*

**OTHER:**

1. *General comment section*

2. *Paperwork* – for instance, if the parents requested results, type of questionnaire administered, etc.
Section C. Additional tables showing the frequency of overarching constructs and their components

Table S1 presents frequency data for several constructs: the target child’s demographic information, demographics of their siblings, information about the interviewer/administrator, the target child’s (pre)school information, the demographics of target child’s parents/caregivers/informants, as well as demographics of any other individuals inquired about.

Table S1. Demographic variables documented across 48 questionnaires (number and percentage of questionnaires documenting each component)

<table>
<thead>
<tr>
<th>Overarching construct (%)</th>
<th>Components measured</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics: child (94%)</td>
<td>Child's name</td>
<td>26</td>
<td>54%</td>
</tr>
<tr>
<td></td>
<td>ID</td>
<td>18</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>18</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>Today's date</td>
<td>27</td>
<td>56%</td>
</tr>
<tr>
<td></td>
<td>Date of birth</td>
<td>30</td>
<td>63%</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>14</td>
<td>29%</td>
</tr>
<tr>
<td></td>
<td>Birth order</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Age/date of arrival</td>
<td>14</td>
<td>29%</td>
</tr>
<tr>
<td></td>
<td>Place of birth</td>
<td>20</td>
<td>42%</td>
</tr>
<tr>
<td></td>
<td>Current residence (location and/or length)</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>Previous residence (location and/or length)</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Nationality</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Ethnicity</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>(Pre)School information (38%)</td>
<td>School</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Day care/Babysitter/Preschool</td>
<td>6</td>
<td>13%</td>
</tr>
</tbody>
</table>
### Teacher
- Teacher: 6 (13%)

### SLT
- SLT: 1 (2%)

### Grade/Class
- Grade/Class: 14 (29%)

### Demographics: siblings (42%)

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td>Number</td>
<td>10</td>
<td>21%</td>
</tr>
<tr>
<td>Sex</td>
<td>10</td>
<td>21%</td>
</tr>
<tr>
<td>Date of birth</td>
<td>6</td>
<td>13%</td>
</tr>
<tr>
<td>Age</td>
<td>13</td>
<td>27%</td>
</tr>
<tr>
<td>Birth order</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td>Place of birth</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Age of exposure</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Languages</td>
<td>2</td>
<td>4%</td>
</tr>
</tbody>
</table>

### Demographics: parents/caregivers/informant (75%)

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td>Relationship to child</td>
<td>21</td>
<td>44%</td>
</tr>
<tr>
<td>Sex</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td>Contact</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td>Age/Date of birth</td>
<td>6</td>
<td>13%</td>
</tr>
<tr>
<td>Place of birth</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td>Country of origin</td>
<td>6</td>
<td>13%</td>
</tr>
<tr>
<td>Current residence</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td>Previous residence</td>
<td>3</td>
<td>6%</td>
</tr>
<tr>
<td>Length of residence in the HL country</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Length of residence in the SL country</td>
<td>8</td>
<td>17%</td>
</tr>
<tr>
<td>Intention to stay in the SL country</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Nationality</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Demographics: other people/generic (10%)</td>
<td>Components measured</td>
<td>n</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------------------</td>
<td>----</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Race</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Languages</td>
<td>18</td>
<td>38%</td>
</tr>
<tr>
<td>Sex</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Age/Date of birth</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Place of birth</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Country of origin</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Current residence</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Length of residence in the SL country</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Interviewer's/Administrator's data (25%)</td>
<td>Components measured</td>
<td>n</td>
</tr>
<tr>
<td>Name</td>
<td>12</td>
<td>25%</td>
</tr>
<tr>
<td>Contact</td>
<td>6</td>
<td>13%</td>
</tr>
<tr>
<td>Who is completing the questionnaire</td>
<td>1</td>
<td>2%</td>
</tr>
</tbody>
</table>

Table S2 presents the frequency of components which could be indicative of the target child’s type of bilingualism (e.g., simultaneous vs. sequential, language dominance).

**Table S2.** *Type of bilingualism variables documented across 48 questionnaires (number and percentage of questionnaires documenting each component)*

<table>
<thead>
<tr>
<th>Overarching construct (%)</th>
<th>Components measured</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of bilingualism (child) (60%)</td>
<td>List of languages</td>
<td>19</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>Stronger language</td>
<td>13</td>
<td>27%</td>
</tr>
<tr>
<td></td>
<td>Type of exposure*</td>
<td>8</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>Date/Age of exposure/acquisition</td>
<td>14</td>
<td>29%</td>
</tr>
</tbody>
</table>

*Note. *This can include: parental strategies regarding which language(s) they use with the child (e.g., one parent, one language), providers/contexts of first exposure, place where the child got into contact, started learning, or started getting exposed to a specific language.*
Information on the language skills of two main caregivers, as well as of other interlocutors is presented in Table S3. Rather than using labels *mother* and *father*, we used *caregiver 1* and *caregiver 2* to cover broader constellations of parents/caregivers. For consistency, whenever a questionnaire inquired about the mother, we classified this information under caregiver 1, while the father-related questions were listed under caregiver 2.

Table S3. Variables related to the proficiency/quality of interlocutors documented across 48 questionnaires (number and percentage of questionnaires documenting each component)

<table>
<thead>
<tr>
<th>Overarching construct (%)</th>
<th>Components measured</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caregiver 1's HL skills/quality (29%)</td>
<td>Listening/ Comprehension</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Speaking/ Production</td>
<td>13</td>
<td>27%</td>
</tr>
<tr>
<td></td>
<td>Overall proficiency</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Errors</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Accent</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Caregiver 2's HL skills/quality (25%)</td>
<td>Listening/ Comprehension</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Speaking/ Production</td>
<td>11</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td>Overall proficiency</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Errors</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Accent</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Additional caregiver's HL skills/quality (2%)</td>
<td>Listening/ Comprehension</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Speaking/ Production</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Overall proficiency</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Caregiver 1’s SL skills/quality (42%)</td>
<td>Speaking/ Production</td>
<td>18</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>Listening/ Comprehension</td>
<td>9</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>Overall proficiency</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>6</td>
<td>13%</td>
</tr>
</tbody>
</table>
Table S4 illustrates the frequency of components which embody the socioeconomic status (SES) of the target child and their family. Holidays are listed when this information was used to estimate the family’s SES. In Table 4 of the paper, some other operationalisations of holiday-related activities are listed under ‘Travel/Holidays’ when this information was not used as an indicator of SES.
Table S4. **Variables related to the socioeconomic status documented across 48 questionnaires (number and percentage of questionnaires documenting each component)**

<table>
<thead>
<tr>
<th>Overarching construct (%)</th>
<th>Components measured</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Socioeconomic status (58%)</strong></td>
<td>Parent/Caregiver/Informant education</td>
<td>20</td>
<td>42%</td>
</tr>
<tr>
<td></td>
<td>Parent/Caregiver/Informant occupation</td>
<td>19</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>Relatives' education and profession</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Free/Reduced meals</td>
<td>3</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td>Property</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Income</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Holidays</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Status</td>
<td>1</td>
<td>2%</td>
</tr>
</tbody>
</table>

Table S5 lists any information on the target child’s early milestones, general behaviour, developmental issues (regarding the child or their family).

Table S5. **Variables related to the development and behaviour status documented across 48 questionnaires (number and percentage of questionnaires documenting each component)**

<table>
<thead>
<tr>
<th>Overarching construct (%)</th>
<th>Components measured</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Developmental issues/ concerns (56%)</strong></td>
<td>Hearing problems</td>
<td>11</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td>Birth/ Pregnancy complications</td>
<td>9</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>Developmental delay</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Language problems</td>
<td>20</td>
<td>42%</td>
</tr>
<tr>
<td></td>
<td>Academic/ Learning problems</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Behavioural/ Social problems</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Physical problems</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Health problems</td>
<td>9</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td><strong>Family history (17%)</strong></td>
<td>(Mainly) Language difficulties</td>
<td>8</td>
<td>17%</td>
</tr>
<tr>
<td><strong>Early milestones (25%)</strong></td>
<td>Beginning to walk</td>
<td>3</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td>Sitting</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Babbling and first words</td>
<td>11</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td>First short sentences</td>
<td>8</td>
<td>17%</td>
</tr>
<tr>
<td>Overarching construct (%)</td>
<td>Components measured</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------------------------------------------------------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Attitudes towards languages/Importance of languages (38%)</td>
<td>Parent/ Caregiver encouraging child to learn</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Beliefs on bi/multilingualism</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>Other beliefs about language/learning</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Child's goals and expectations</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Parental goals and expectations</td>
<td>10</td>
<td>21%</td>
</tr>
<tr>
<td></td>
<td>Preferred language (child)</td>
<td>9</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>Preferred language (parents)</td>
<td>3</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table S6 presents any data relating to attitudes towards the languages of bilinguals, either by parents/caregivers or by children themselves.

Table S6. Variables related to attitudes documented across 48 questionnaires (number and percentage of questionnaires documenting each component)

Finally, Table S7 lists a miscellaneous set of components not classified under any constructs above. These include components relating to language loss, as well as general/technical questions that sometimes appear in questionnaires quantifying bilingual experience.
<table>
<thead>
<tr>
<th>Overarching construct (%)</th>
<th>Components measured</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition (4%)</td>
<td>Language loss (child)</td>
<td>2</td>
<td>4%</td>
</tr>
<tr>
<td>Other (17%)</td>
<td>General comment section</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Paperwork</td>
<td>2</td>
<td>4%</td>
</tr>
</tbody>
</table>
Section D. A list of questionnaires identified after the data collection and the manuscript submission

The questionnaires listed in this section were identified after the data collection and the initial manuscript submission. They were collected in one of the following ways: (1) emailed to us by the authors, or (2) while reading literature on bilingualism. We list them here as the readers might find them useful resources.

1. CECER-DLL Child and Family and Teacher Questionnaires (Castro et al., 2020)
2. Contextual and Individual Linguistic Diversity Questionnaire (CILD-Q) (Wigdorowitz et al., 2020)
3. Language and Social Background Questionnaire, Version for Children (Anderson et al., 2020)
4. Language and Social Background Questionnaire, Version for Older Adults (Anderson et al., 2020)
References:


De Houwer, A. (2016a). *PEGEBOS-1. Polish-German and English-German Early Bilingualism Online Survey Wave 1 @ child age 2;0*. [Unpublished manuscript]. Retrieved from the author.


https://doi.org/10.1177/1367006917733066


