

This is an author version of the article published as:

Picture-based vocabulary assessment versus parental questionnaires: A cross-linguistic study of bilingual assessment methods

Pernille Hansen, Magdalena Łuniewska, Hanne Gram Simonsen, Ewa Haman, Karolina Mieszkowska, Joanna Kołak, Zofia Wodniecka

International Journal of Bilingualism, First Published October 12, 2017,
<https://doi.org/10.1177/1367006917733067>

Do indirect and direct measures of children's language skills correspond? A comparative study of cross-linguistic language assessment

Pernille Hansen¹, Magdalena Łuniewska^{2,3}, Hanne Gram Simonsen¹, Ewa Haman², Karolina Mieszkowska², Joanna Kołak^{4,2}, Zofia Wodniecka⁵

¹ Center for Multilingualism in Society across the Lifespan (MultiLing), University of Oslo

² Faculty of Psychology, University of Warsaw

³ Nencki Institute of Experimental Biology, Polish Academy of Sciences

⁴ International Centre for Language and Communicative Development (LuCiD), University of Manchester

⁵ Institute of Psychology, Jagiellonian University

Corresponding author:

Pernille Hansen, Center for Multilingualism in Society across the Lifespan (MultiLing), University of Oslo, P.O. Box 1102 Blindern, 0317 Oslo, Norway.

Email: pernille.hansen@iln.uio.no

Abstract

Purpose: As a contribution to the complex task of developing appropriate tools for language assessment in bilingual children, the current paper investigates whether direct and indirect measures of language skills paint a similar picture of children's multilingual language competence.

Approach: Emerging from the recent COST Action IS0804 (Bi-SLI), the study used two new tools from the battery Language Impairment Testing in Multilingual Settings (LITMUS): the direct assessment tool Cross-linguistic Lexical Tasks (CLT) and the parental report Parents of Bilingual Children Questionnaire (PaBiQ), offering an indirect measure of overall language skills.

Data and Analysis: The participants were 36 children of Polish immigrants to Norway or the UK.

Correlations were investigated with Kendall's rank correlation, and group comparisons carried out with Wilcoxon rank sum tests.

Findings: The direct and indirect measures correlated. There were no group differences on the direct measure, but the parents in the UK still judged their children as less proficient in Polish than the parents in Norway did. Two different accounts for this incongruity are discussed: First, parents in the UK may set higher benchmarks for their children's minority language skills than the parents in Norway. Alternative accounts of this interpretation related to differences in the parents' socio-economic background, proficiency or language attitudes are discussed. Second, parental reports may indicate early stages of attrition of the minority language among the children in the UK that the direct lexical assessment tool may not be sensitive enough to uncover.

Originality: The study used two new tools designed specifically for multilingual children – comparing two groups of children of a recent and growing immigration group, whose language development is currently underinvestigated.

Implications: The findings underscore the complexity of assessing bilingual children's full language competence. The cross-cultural differences documented here call for further longitudinal research systematically comparing immigrant children from different language backgrounds.

Keywords: language assessment, language acquisition, migration, Cross-linguistic Lexical Tasks, PaBiQ, immigrant children, language shift, Polish

Introduction

In the wake of the 2004 EU enlargement, many Western European countries have seen a rapid increase in the immigration from the former Eastern Bloc. As a result, Poland is currently the most common country of birth in the immigrant population of both Norway (Statistics Norway, 2016) and the UK (Office for National Statistics, 2016). Migrant communities may either retain their own minority language(s) at the expense of the majority language(s), maintain their minority language(s) alongside the majority language(s), or undergo a process of language shift, where the heritage language is replaced with the majority language (Fishman, 1991). As pointed out by Fishman (1991) the first two outcomes depend on intergenerational transmission of the minority language; language transmission in turn depends on language maintenance. Language shifts may vary in speed within different contexts (Gal, 1979), but they tend to progress from one generation to the next (De Houwer, 2007; Fishman, 1991; Saltarelli & Gonzo, 1977).

Importantly, an individual can become or cease to be multilingual; as such, multilingualism must be seen as a dynamic state rather than a static property (Grosjean, 2008). Thus, an individual may acquire a minority language in a monolingual setting from birth, become multilingual as a pre-schooler and be a monolingual speaker of the majority language as an adult. Multilingualism is not only dynamic, but also multifaceted: A multilingual speaker's languages may have complementary functions and usage patterns. Thus, to fully capture multilingual children's language development, one must study all their languages (De Houwer, 2009; Grosjean, 2008; Pearson, 2010). Even so, it is common to assess only the majority language, in research (e.g. Bialystok, Craik, & Luk, 2008; Lervåg & Aukrust, 2010; Melby-Lervåg & Lervåg, 2013) as well as in speech-language pathology (Bedore & Pena, 2008; Paradis, Emmerzael, & Duncan, 2010).

In clinical settings, this monolingual practice means multilingual children risk being misdiagnosed with language impairment (Armon-Lotem & de Jong, 2015; Bedore & Pena, 2008; de Jong, Çavuş, & Baker, 2010; Kohnert, 2010; Leonard, 2014; Paradis, 2016). Children with a minority home language may for instance have smaller vocabularies in the majority language than their age-matched monolingual peers (Bialystok et al., 2008; Melby-Lervåg & Lervåg, 2013; Oller & Eilers, 2002), similarly to monolingual children with a language impairment (Leonard, 2014; Leonard & Deevy, 2004; Paradis, 2010; Rice & Hoffman, 2015). Importantly, whereas slow development only in the majority language may be attributable to the language environment (Paradis et al., 2010), multilingual children with language impairment will show atypical patterns across all of their languages (Kohnert, 2010; Paradis, 2016).

Thus, cross-linguistic language assessment is crucial for the identification of language impairment in multilingual populations. However, there is a lack of appropriate tools for the combinations of languages spoken by immigrant populations (Peña, 2007). Direct assessment tools that are available across multiple languages are typically direct translations from the English version. For instance, the Norwegian version (Lyster, Horn, & Rygvold, 2010) of British Picture Vocabulary Scale (BPVS) is largely, a direct translation of its British English counterpart (Dunn & Dunn, 2009). This method poses challenges to the validity of cross-linguistic comparisons because the 'same' items may not be equally difficult across languages (Peña, 2007).

Peña (2007) argues that to ensure equivalence across languages, assessment tool construction must consider item difficulty, by including measures such as words' frequency of occurrence in the target language or their age of acquisition (AoA). Importantly, an assessment tool for children may need to rely on other measures than a tool for adults; as demonstrated by Goodman, Dale, and Li (2008) and Hansen (in press), frequency in child-directed speech is a better predictor of when children acquire words than frequency in adult written language.

A recent development in this respect is the assessment tool battery Language Impairment Testing in Multilingual Settings (LITMUS) (Armon-Lotem, de Jong, & Meir, 2015), developed through the recent European network COST Action IS0804 *Language impairment in a multilingual society: Linguistic patterns and the road to assessment* (2009-2013). One of the new tools developed by members of this network is *Cross-linguistic Lexical Tasks* (CLT) (Haman, Łuniewska, & Pomiechowska, 2015), a tool that has incorporated AoA in its construction procedure, as suggested by Peña (2007).

Several studies of CLT results have documented that target words acquired early in life according to this measure, are easier for monolingual as well as bilingual children than target words with a high AoA (Altman, Goldstein, & Armon-Lotem, in press; Haman et al., in press; Hansen, Simonsen, Łuniewska, & Haman, in press). Investigating CLT results from monolingual Polish and Norwegian children as well as bilingual Polish-Norwegian children, Hansen et al. (in press) reported that AoA accounted better for the difficulty of the CLT target words in both languages and among both groups than did the frequency of occurrence in CDS. They found no difference in performance between the Polish and Norwegian monolinguals, and argued that for these two languages at least, CLT succeeds to be cross-linguistically equivalent (Hansen et al., in press).

Another issue for a valid and reliable assessment across the languages of multilingual children concerns language competence: Educators and speech-language pathologists rarely understand the minority languages spoken by the children in their care (Williams & McLeod, 2012). Given reliable tools and rigid methodologies, a valid cross-linguistic assessment may be carried out with help from interpreters. Computer-based tools with pre-recorded audible stimuli and automatic recording and scoring of the results may further aid the replicability of the assessment, as this restricts the influence of the individual experimenter.

However, even with adequate tools, assessing children's skills in their home language will still be a challenge for practitioners that are not proficient in the language. Another solution is indirect assessment through parental reports. Combining different adaptations of the MacArthur-Bates CDI (Fenson et al., 2007) may be a valid option for children up to age 3 (Conboy & Thal, 2006; De Houwer, Bornstein, & Putnick, 2014; Elin Thordardottir, 2015; Gatt, Grech, & Dodd, 2015; Law & Roy, 2008; O'Toole et al., 2016; Pearson, Fernandez, & Oller, 1993). For children up to age 7, Restrepo (1998) and Paradis et al. (2010) have demonstrated that parents' judgments about skills in the home language can discriminate between language impairment and typical language development of early second language learners of English in the US or Canada.

Although parents are generally good judges of their children's language skills (Fenson et al., 1994), they may misjudge their children's skills in the majority language, particularly if they themselves are new to the language (Tuller, 2015). Furthermore, both parental and societal expectations may affect where parents set the benchmark. Hence, similar parental judgments across different language communities do not necessarily entail similar levels of language skills, and vice versa.

Research questions

This study investigates the language skills of immigrant children in Norway and the UK acquiring Polish at home, combining the lexical assessment tool CLT (Haman et al., 2015) and the Parents of Bilingual Children Questionnaire (PaBiQ) (COST Action IS0804, 2011; Tuller, 2015), in part based on the *Alberta Language Environment Questionnaire* (Paradis, 2011) and the *Alberta Language and Development Questionnaire* (Paradis et al., 2010). The research questions are:

1. Do indirect and direct measures of immigrant children's language skills correlate within and across language communities?
2. Do the two types of measurement paint a similar picture of the children's language competence?

Methods

The current paper compares indirect and direct measures of language skills among children of recent Polish immigrants to the UK or Norway. The participant groups were chosen for three reasons: First, recent Polish immigrants are numerous in both countries. Second, both countries offer affordable childcare and have a high formal childcare coverage rate (Eurostat, 2016; Mills et al., 2014). Third, the two languages acquired by these children (Polish vs English/Norwegian) are not closely related, whereas the two majority languages (English and Norwegian) are, making the task linguistically comparable.

To provide indirect measures of the child's current language skills, as well as the linguistic, developmental and socio-economic background, the parents were asked to fill in on paper a Polish pilot version of the PaBiQ (COST Action IS0804, 2011; Tuller, 2015), *Kwestionariusz Rozwoju Językowego (KRJ)* [Questionnaire on Language Development] (Kuś, Otwinowska, Banasik, & Kiebzak-Mandera, 2012). For the direct language assessment, lexical skills were measured with the Cross-linguistic lexical tasks (CLT) (Haman et al., 2015) a lexical assessment tool designed for multilingual children. The participants and the two tools are presented below.

Participants

The participants are 36 children (aged 4;2 – 6;6) of Polish immigrants to the UK or Norway, living either in the region around the capital (London/Oslo) or a relatively large city (Aberdeen/Bergen) (see Table 1). The families were recruited through day-care facilities, schools, Polish newspapers in Norway and the UK, speech therapists, communities at universities, portals about parenting, community groups in social media, Polish Saturday Schools, catholic churches and Polish shops (for a discussion of the recruitment, see Haman, Wodniecka, Kołak, Łuniewska, & Mieszkowska, 2014). Presumably, all the children were typically

developing; none had been referred to a speech language therapist, and none were at a high risk of language impairment, according to information on early linguistic milestones, parental concern regarding language development, and the history of language difficulties in the family (see Tuller et al., 2013).

Table 1. Distribution of age and gender among the two participant groups.

Country	n	Age mdn (range)	F:M
Norway	18	5;3 (4;3–5;10)	5:13
UK	18	5;8 (4;2–6;6)	7:11

According to information from the background questionnaire, all the children lived with both parents, all of whom were L1 speakers of Polish. Whereas 20 of the 36 mothers had higher education, the same was true for only nine of the fathers (see also Haman et al., 2014). According to Fisher’s exact tests, the proportions of highly educated mothers did not differ significantly between the two groups (Norway: 9, the UK: 11, $p=0.738$), whereas there were significantly more fathers in the UK (8) than in Norway (1) with higher education ($p=0.018$). Two mothers and six fathers had only basic education, all of whom resided in Norway. These group differences correspond to differences in the populations: Whereas “the most recent migration to English-speaking countries is the domain of young and relatively well-educated persons” (Kaczmarczyk, 2010, p. 175), Polish immigrants to Norway tend to be slightly older and not have higher education (Friberg, 2012).

Ethical considerations. The assessment in Norway was approved by the Norwegian Social Science Data Services, and the assessment in the UK was approved by the Ethics Committee at the Faculty of Psychology, University of Warsaw, Poland. The parents were duly informed about the study, and they signed a written consent. The children received a small gift for their participation, whereas neither parents nor day-care facilities received any recompense. The parents were informed that they would not get diagnostic feedback of an individual child performance; the tools used here aim to help identify language impairment in multilingual children, but they are not ready for clinical use as norms have not yet been established. Thus, only analyses on group level are possible at present.

Parental questionnaire

In the current study, the questionnaire data are utilised for three different purposes: First, to exclude children with a high risk of language impairment (see above), second, to portray the participants, and third, to provide an indirect assessment of the children’s current linguistic skills across their languages for the comparison with CLT results. Concerning the second point, participants were profiled by five factors derived from the questionnaire data, following Tuller (2015) and Łuniewska, Kołak, and Kacprzak (unpublished manuscript): The age of onset of exposure to each language, three measures of current language exposure and use, and finally, the parents’ self-rated proficiency in the majority language. Regarding the third point, the parents rated their child’s skills in each language, covering phonology, vocabulary, syntax and general communicative skills.

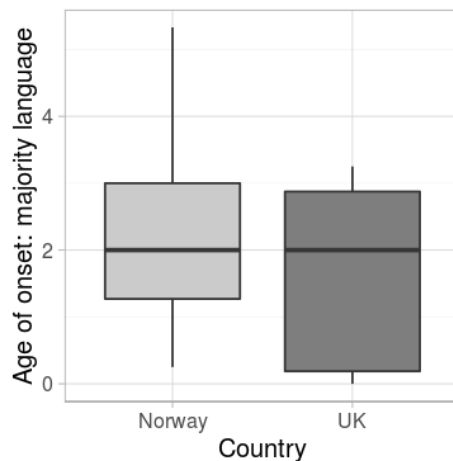


Figure 1. Age of onset of exposure to the majority language, divided by participant group.

Age of onset of language exposure. According to the background questionnaire, all children had been exposed to Polish from birth. The median child had been exposed to the majority language since age 2;0, but there was considerable diversity among the participants: Five participants had reportedly heard the majority language from birth, and another four had been exposed to the majority language already before their first birthday. Note that these nine children (three in Norway and six in the UK) had two Polish parents, and reportedly heard and used Polish more than the majority language with other family members. On the other end of the scale, one of the participants was 4;6 upon first contact with the majority language, Norwegian. Figure 1 illustrates the distributions of age of onset of exposure to the majority language among the two participant groups; according to a Wilcoxon rank sum test, the difference between the two groups is not significant ($W=189, p=0.401$).

Depending on definitions, these children's acquisition of the majority language could be considered either as simultaneous with the minority language, because exposure started before age five (Meisel, 2004), or as early second language acquisition (De Houwer, 2009), as all parents reported to mainly speak Polish to their children.. Based on the age of onset alone, one may argue that some experienced bilingual first language acquisition (De Houwer, 2009), but this is contradicted by the low amount of majority language input. The goal of the current paper, to compare indirect and direct assessments of language skills, does not at any rate rest upon the categorisation. However, in order to improve the recognition of language impairment in multilingual children, it is important to capture the variation in the population.

Current language exposure and use. In the background questionnaire, parents report the patterns of language use in the home by specifying how often ('never', 'seldom', 'sometimes', 'often' and 'always') each language is used to the child and by the child in conversations with each parent, sibling and other caregiver living with the family (e.g. grandparents). These data were used to gauge the language balance in the child's home input and home output. Estimated frequencies of use were weighted giving 0 points if a language was 'never' used by a speaker, and 4 points if it was 'always' used, giving twice the weight to parents and siblings as to other caregivers (Łuniewska et al., unpublished manuscript). The patterns of

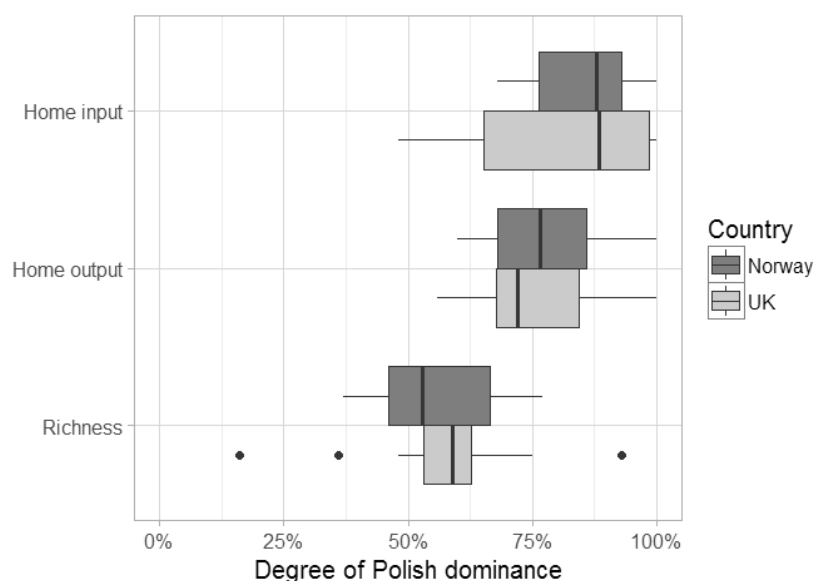


Figure 2. Degree of Polish dominance in the input from and output to other family members

language use in the family may be key to early language development, but the maintenance of a minority language also rests upon interaction with friends and acquaintances, and use in activities such as singing and reading (Fishman, 1991). The current study employs a measure of language richness that incorporates language use with friends and friends of the family and in a set of leisure activities (reading, watching TV or movies, computer activity and children’s songs or nursery rhymes), following Tuller (2015).

The scores for each language were used to estimate the degree of Polish dominance in the input to and output from each child, as well as their language richness, illustrated in Figure 2. As evident from Figure 2, the majority of the children are biased towards Polish in their input as well as their output, although slightly less so in the latter, whereas the measures of language richness are more balanced between the languages. The two groups do not differ significantly in input ($W=166, p=0.924$), output ($W=184, p=0.494$) or richness ($W=145, p=0.601$), but there is a strikingly large variation in the input of the UK participants, compared to those in Norway.

Parents’ self-evaluation of majority language proficiency. In PaBiQ, parents rate their own proficiency in each of the languages they know on a five-point scale from ‘only a few words’ to ‘very well’. While as many as 12 of 18 participants in Norway have at least one parent reporting less than basic abilities in the majority language, the same is true for only four of the 18 UK participants. The difference is significant, according to a Fisher’s exact test ($p=0.018$).

Parental judgments of children’s overall language skills. The pilot version of PaBiQ used here asks parents nine questions about their child’s current skills in their languages, each rated on a four-point scale. The questions cover phonology, vocabulary, syntax and general communicative skills, and tap into both expressive and receptive language (see Appendix 1). The current paper follows Tuller (2015) in using the parental judgment to calculate a sub-score of the child’s overall language skills, ranging from 0 (lowest

possible rating on all questions) to 27 (highest possible rating on all nine questions); the judgments for the participants ranged from 3 to 27, with a median of 18 points. In addition, the question regarding lexicon size was used to divide the children into two groups: those with reportedly smaller vocabulary than their peers and those with vocabulary size reportedly similar to their peers.

These scores are used for two purposes. First, this indirect measure of overall language skills is compared to direct measures of lexical tasks, collected by means of the lexical assessment tool CLT. Second, the study investigates whether the direct and indirect measures agree in their evaluation of the participants.

The Cross-linguistic Lexical Tasks

Direct assessment of the children's lexical skills was carried out with the CLT in Polish (Haman, Łuniewska, Pomiechowska, Szewczyk, & Wodniecka, 2012), UK English (Haman, Łuniewska, Polinska, & Mieszkowska, 2012) and Norwegian (Simonsen, Hansen, & Łuniewska, 2012). The tool consists of four parts: comprehension and production of nouns and verbs. In the comprehension tasks, the participants hear a target word and choose between four different pictures shown simultaneously on the screen. In the production tasks, they name a single depicted object or action. Each task has 32 items. The various language versions of CLT have been designed independently of each other, based on a strict procedure devised to ensure cross-linguistic comparability (Haman et al., 2015). Recent studies have indicated that the tool yields comparable results across a wide variety of languages (Haman et al., in press; Hansen et al., in press).

CLT assessment and scoring. All children were assessed in a quiet room in their day-care facility with the computer version of CLT (e-CLT), where questions are pre-recorded and played automatically. Generally, the gap between assessment in each language was about a week. L1 speakers carried out the assessment of Polish (in both countries) and Norwegian (in Norway), whereas in the UK, highly proficient L2 speakers of English residing in the UK conducted the English assessment. During the assessment, the experimenter only addressed the child in the tested language.

Each of the four task sets (comprehension and production of nouns and verbs) consists of 32 items in a fixed order, but the order of the four subtasks and the order of the two languages assessed were counterbalanced across children. Each session (four tasks per language) lasted about 15 minutes. The children received age-appropriate information about the testing, and they were told that they could terminate at any time. All the UK participants completed the tasks in both languages, but three of the participants in Norway did not complete the Polish assessment, and four did not complete the Norwegian assessment. For these seven, only data from the completed language are included in the analyses.

E-CLT automatically scores comprehension responses, and records production responses for manual transcription and scoring. In this study, any response involving the root of the target word in the assessed language was considered correct, along with regional variants and synonyms. Polish-English bilinguals coded the UK production data, whereas the Norwegian data were coded through joint efforts between Polish and Norwegian researchers. The first and second author (native speakers of Norwegian and Polish respectively) carefully checked the data from Norwegian together to ascertain consistent coding, and to

recognize and appropriately score responses involving both languages. The CLT results from each language may potentially range from 0 (no correct answers) to 128 (correct answers on all items); the participants' scores ranged from 33 to 120, with a median of 87.

Analyses

All statistical analyses were carried out in R (R Core Team, 2015) using RStudio (Team, 2015). Both measures analysed here are skewed towards top scores; Shapiro-Wilk normality tests revealed significant divergence from a normal distribution for both parental judgments ($W=0.96, p=0.022$) and CLT results ($W=0.94, p=0.003$). The study hence focuses on the order of the participants rather than the scores: Correlations between CLT results and parental judgments of children's skills are investigated with Kendall's rank correlation tau (τ), and group comparisons are carried out with Wilcoxon rank sum tests. Group comparisons are carried out both within the same language across countries and within the same countries across languages, with p values adjusted with Holm correction.

Results

Figure 3 illustrates the relationship between CLT results and parental judgments. Overall, there is a significant correlation between the indirect assessment, the parental judgments of overall language skills, and the direct assessment, the children's CLT results ($r_\tau=0.44, p<0.001$). The two measures correlate also within the majority language results (UK English in the UK and Norwegian in Norway) ($r_\tau=0.39, p=0.002$), but not within the Polish results from the two groups ($r_\tau=0.14, p=0.28$). This lack of correlation is observable

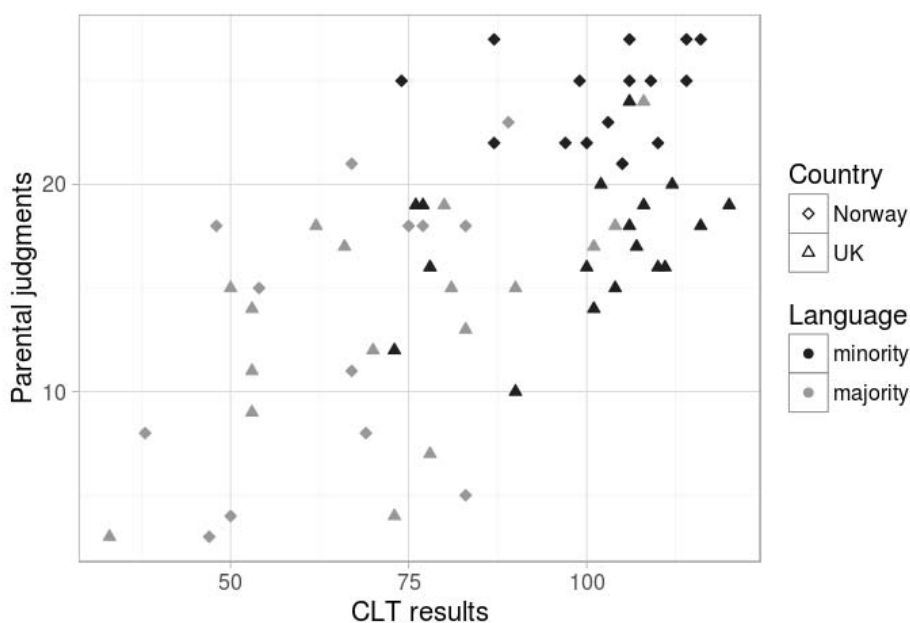


Figure 3. CLT results as a function of parental judgments, by language and country.

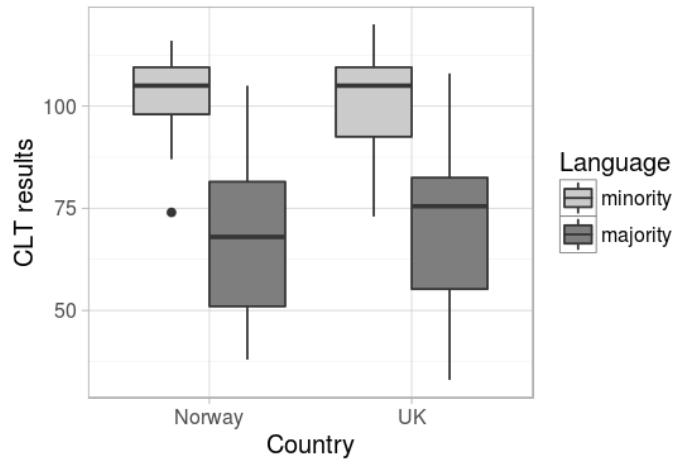


Figure 4. Boxplot of CLT results by country and language.

in Figure 3: most children from both groups score high on the Polish CLT. However, the UK parents judge their children's skills as lower than the parents living in Norway do, even if the CLT scores are comparable.

The group similarities among CLT results are even more apparent in Figure 4. The Polish scores surpass the scores in the majority language for both the British ($W=273$, $p=0.001$) and the Norwegian group ($W=196$, $p<0.001$), and there are no significant between-group differences in neither Polish ($W=137$, $p=0.971$) nor the majority language ($W=107$, $p=0.940$).

Figure 5 shows the parental judgments of overall language skills, telling a slightly different story: The children in Norway have significantly higher skills in Polish than in Norwegian, according to their parents ($W=300$, $p<0.001$), but for the UK group, there is no significant language difference ($W=233$, $p=0.050$). The parental judgments of the children's skills in the majority language are not significantly different across the two countries ($W=159$, $p=0.937$), but the parents residing in the UK judged their children's Polish skills as lower than do the parents residing in Norway ($W=292$, $p<0.001$). Isolating the scores on the indirect assessment of vocabulary size from the indirect compound measure, 25 of the 36 children reportedly knew fewer words in the majority language than their peers, with no significant difference between the two countries, according to Fisher's exact test (UK: 10, Norway: 15, $p=0.15$). Only two children, one from each group, were estimated to know fewer words in Polish than other children of the same age do. These also hold the lowest scores on the overall indirect measure of minority language skills. Thus, it appears that within the parental judgments of vocabulary skills alone, the picture is more similar to the CLT results than to the indirect compound measure: Parents in both groups judge their children to know more words in Polish than in the majority language.

Discussion

The current paper combined the background questionnaire PaBiQ and the lexical assessment tool CLT to investigate Polish-English and Polish-Norwegian children, asking whether there is a correspondence between parental judgments and direct measures of children's language skills. A significant correlation was

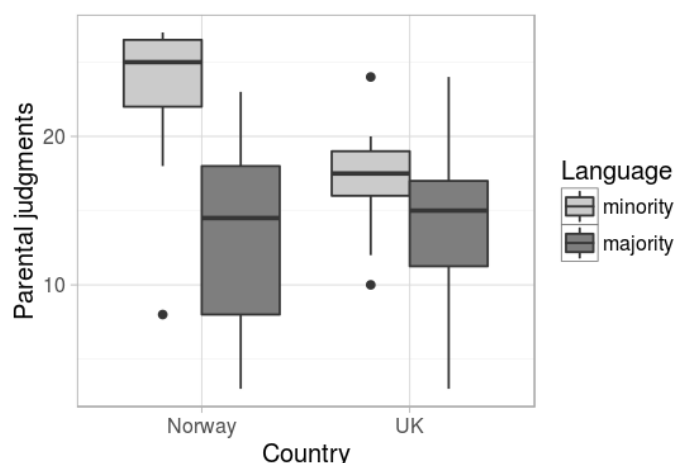


Figure 5. Boxplot of parental judgments of children's skills, by country and language.

found between CLT results and parental judgments, both overall and within the majority language, indicating that the two measures do correspond. No correlation was found between parental judgments and CLT results for Polish. The explanation may lie in the distribution of the results: Although there is a considerable variation, there is a tendency towards a ceiling effect in Polish on both measures (see Figure 3), potentially masking a significant correlation.

The parents residing in the UK systematically judged their children's skills in Polish as lower than the parents in Norway did, even though there was no significant cross-national difference in the children's CLT results in Polish. This incongruity pairs with the surprising lack of a significant correlation between Polish CLT results and parental judgments. There are two possible explanations: First, the differences in parental judgments in the two populations may not be due to differences in skills, but because of distinctive benchmarks set by the parents. As stated above, the parental judgments of children's current language skills are calculated from nine questions (see Appendix) regarding the parent's and the child's satisfaction with the communicative skills, and whether the child is able to hold a conversation in each language. These are context-dependent questions, and the answers may rest upon parental language ideologies, affected by political, cultural and economic factors (Curdt-Christiansen, 2009).

Alternatively, CLT may fail to reveal differences in language performance that the parents are sensitive to. Whereas CLT only measures lexical skills, and includes only early-acquired nouns and verbs, the parents may tap into observations of their children's morphological, pragmatic or interactional skills. If we rest upon the parental judgments, the UK group is more balanced between their languages than the Norwegian group. Although all the children (in both groups) used Polish more than the majority language at home, conforming to De Houwer (2007), they used the majority language more than other family members did when conversing. Thus, the participants could be on their way towards a language shift from the minority to the majority language, with the UK children further along the path than their peers in Norway. Alternatively, they may be on the path of harmonious bilingual development, about to reach equal levels of proficiency in each language (De Houwer, 2013).

According to the questionnaire, there were no significant group differences in the families' language practices. However, the parents in the UK were reportedly significantly more proficient in English than the parents in Norway were in Norwegian. This difference is unsurprising, as English is a global language taught in Polish schools, while few learn Norwegian before moving to Norway. Nevertheless, it means that most of the participants in Norway must depend on Polish when communicating with their parents, whereas intergenerational communication does not stand in the way of a language shift for most of their UK peers. Importantly, although proficiency in the majority language may generally be linked to success and privileges (Lane, 2010), Norwegian does not share the global status of English, which may be viewed by the parents as "an international super language through which a great many social and economic goals can be achieved" (Curdt-Christiansen, 2009, p. 363).

Differences in the education systems may contribute to an earlier language shift within the UK group. The British group were slightly older than the Norwegian group and compulsory education starts one year earlier in the UK. Thus, most of the British participants attended school when they took part in the study, whereas the children in Norway attended full-time formal childcare. That is, both groups spent their days in an institution where the majority language was the primary language, but possibly with more emphasis on majority language teaching in the case of the UK participants.

Limitations and future directions

The current study is limited by the available data, and first and foremost by the number of participants. Recruiting participants proved to be more difficult than foreseen, even with the multitude of channels used to reach Polish families in the two countries (see the methods section, and Haman et al., 2014). The limited number of participants calls for caution regarding statistical methods; the questionnaire offers information on a variety of factors that may affect children's performance on a lexical test or parental judgments of their children's language skills, but to compare the potential effects of these factors, more data are needed. One possible direction for future studies would be to include data from other groups of bi- and multilinguals; this could in addition shed further light on the comparability of the two tools.

A caveat to this study is that with two incongruent measures of language skills, we cannot determine which of the tools to trust. A third tool could tip the scale. Other teams have combined the tools used here with other tools from the LITMUS battery (Armon-Lotem et al., 2015), and further investigations could resolve whether we should trust the CLT results, indicating that the UK group's Polish skills surpass their English skills, or rather rely on the parental judgments, indicating a balance between the languages.

Conclusion

This paper has documented an overall correlation between parental judgments, measured by the Polish pilot version (Kuś et al., 2012) of the background questionnaire PaBiQ (COST Action IS0804, 2011; Tuller, 2015), and direct measures of Polish-English and Polish-Norwegian children's lexical skills, measured by Cross-linguistic Lexical Tasks (Haman et al., 2015). There was also a significant correlation within the

majority language both within and across language communities. However, within Polish, the parental judgments and CLT results did not correlate, and there was an incongruity between the groups: The CLT results from both the minority and the majority language were comparable across the two countries, but parents residing in the UK judged their children as less proficient in Polish than the parents in Norway did.

The reason for this incongruity could be that the two groups of parents set different benchmarks for their children's minority language skills. However, it is also possible that the UK children are shifting towards the majority language, mediated by the status of the language and the high proficiency among their parents, whereas their peers in Norway, whose parents speak little Norwegian, have (at least so far) maintained their minority language. As the CLT was created to aid the identification of language impairment in multilinguals, and the target words denote concrete objects and actions that presumably are quite frequent in children's lives, the tool may not be sensitive enough to uncover early stages of language attrition. Importantly, to evaluate the possible accounts discussed above, there is need for further research systematically comparing immigrant children from different language backgrounds over time.

Acknowledgements

We wish to thank all the participants, their parents and day-care centre employees for taking part in the study. We are grateful to Elisabeth Holm, Katarzyna Chyl, Małgorzata M. Haman and Ingeborg Ribbu for collecting and coding parts of the data, and to Ewa Wapinska and Emilia Dymarczyk for aiding the collection of Polish data in Norway. We would also like to thank members of COST Action IS0804 for discussions on the development of CLT and preliminary results from this study, and Pia Lane and Bente Ailin Svendsen for helpful comments to the manuscript.

Funding

This study was partially supported by the following institutions: National Science Centre (Poland; grant no. 809/N-COST/2010/0); Polish Ministry of Science and Higher Education (contract no. 0046/DIA/2013/42); Council of Norway through its Centres of Excellence funding scheme (project number 223265); and the European Cooperation in Science and Technology (COST) through its funding of Short-Term Scientific Missions (ECOST-STSM-IS0804-121112-023461).

References

- Altman, C., Goldstein, T., & Armon-Lotem, S. (in press). Quantitative and qualitative differences in the lexical knowledge of monolingual and bilingual children. *Clinical Linguistics and Phonetics*.
- Armon-Lotem, S., & de Jong, J. (2015). Introduction. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: disentangling bilingualism from Language Impairment* (pp. 1-22). Bristol: Multilingual Matters.
- Armon-Lotem, S., de Jong, J., & Meir, N. (Eds.). (2015). *Methods for assessing multilingual children: disentangling bilingualism from Language Impairment*. Bristol: Multilingual Matters.
- Bedore, L. M., & Pena, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11(1), 1-29.
- Bialystok, E., Craik, F. I., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, 21(6), 522-538.
- Conboy, B. T., & Thal, D. J. (2006). Ties Between the Lexicon and Grammar: Cross - Sectional and Longitudinal Studies of Bilingual Toddlers. *Child Development*, 77(3), 712-735.
- COST Action IS0804. (2011). *Parents of Bilingual children Questionnaire(PaBiQ)*. A part of the LITMUS Battery (COST IS0804). Retrieved from <http://www.bi-sli.org>
- Curdt-Christiansen, X. L. (2009). Invisible and visible language planning: Ideological factors in the family language policy of Chinese immigrant families in Quebec. *Language Policy*, 8(4), 351-375.
- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied psycholinguistics*, 28(03), 411-424.
- De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol: Multilingual Matters.
- De Houwer, A. (2013). Harmonious bilingual development: Young families' well-being in language contact situations. *International Journal of Bilingualism*, 1367006913489202.
- De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, 35(06), 1189-1211.
- de Jong, J., Çavuş, N., & Baker, A. (2010). Language impairment in Turkish-Dutch bilingual children. In S. Topbaş & M. Yavaş (Eds.), *Communication disorders in Turkish* Bristol: Multilingual Matters.
- Dunn, L. M., & Dunn, D. M. (2009). *British Picture Vocabulary Scale. Third Edition (BPVS III)*. London, UK: GL Assessment.
- Elin Thordardottir. (2015). Proposed diagnostic procedures for use in bilingual and cross-linguistic contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: disentangling bilingualism from Language Impairment* (pp. 331-358). Bristol: Multilingual Matters.
- Eurostat. (2016). Early childhood and primary education statistics. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Early_childhood_and_primary_education_statistics

- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5), 1-185.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories: User's guide and technical manual* (2 ed.). Baltimore, MD: Paul H. Brookes Publishing Company.
- Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages* (Vol. 76). Clevedon, UK: Multilingual matters.
- Friberg, J. H. (2012). The Stages of Migration. From Going Abroad to Settling Down: Post-Accession Polish Migrant Workers in Norway. *J. of Ethnic and Migration Studies*, 38(10), 1589-1605.
- Gal, S. (1979). *Language shift: Social determinants of linguistic change in bilingual Austria*. New York, NY: Academic Press.
- Gatt, D., Grech, H., & Dodd, B. (2015). Early lexical expression in children exposed to mixed input: A case of monolingual or bilingual development? *International Journal of Bilingualism*, 1367006915572399.
- Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, 35(3), 515-531.
- Grosjean, F. (2008). *Studying bilinguals*. New York, NY: Oxford University Press, USA.
- Haman, E., Łuniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekić, J., . . . Armon-Lotem, S. (in press). Noun and verb knowledge in monolingual preschool children across 17 languages: data from Cross-linguistic Lexical Tasks (CLT). *Clinical Linguistics & Phonetics*.
- Haman, E., Łuniewska, M., Polisenska, K., & Mieszkowska, K. (2012). *Crosslinguistic Lexical Tasks: Polish version (CLT-UK)*. A part of LITMUS Battery (COST IS0804). Faculty of Psychology, University of Warsaw. Unpublished material.
- Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing Cross-linguistic lexical tasks (CLTs) for bilingual preschool children. In S. Armon-Lotem, J. d. Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from Language Impairment* (pp. 196-240). Bristol, UK: Multilingual Matters.
- Haman, E., Łuniewska, M., Pomiechowska, B., Szewczyk, J., & Wodniecka, Z. (2012). *Crosslinguistic Lexical Tasks: Polish version (CLT-PL)*. A part of LITMUS Battery (COST IS0804). Faculty of Psychology, University of Warsaw. Unpublished material.
- Haman, E., Wodniecka, Z., Kołak, J., Łuniewska, M., & Mieszkowska, K. (2014). Social aspects of psycholinguistic research: Reflections on the ongoing study of cognitive and linguistic development of Polish immigrant children. In M. Olpinska-Szkiełko & L. Bertelle (Eds.), *Zweisprachigkeit und Bilingualer Unterricht*. Frankfurt am Main: Peter Lang.

- Hansen, P. (in press). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. Manuscript submitted for publication. *First Language*.
- Hansen, P., Simonsen, H. G., Łuniewska, M., & Haman, E. (in press). Validating the psycholinguistic aspects of CLT: Evidence from Polish and Norwegian. *Clinical Linguistics & Phonetics*.
- Kaczmarczyk, P. (2010). Brains on the move? Recent migration of the highly skilled from Poland and its consequences. In R. Black, G. Engbersen, M. Okólski, & C. Pantîru (Eds.), *A continent moving west? EU Enlargement and Labour Migration from Central and Eastern Europe* (pp. 165-186). Amsterdam, The Netherlands: Amsterdam University Press.
- Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*, 43(6), 456-473.
- Kuś, K., Otwinowska, A., Banasik, N., & Kiebzak-Mandera, D. (2012). *Kwestionariusz Rozwoju Językowego (KRJ) [Questionnaire on Language Development]*. Faculty of psychology, University of Warsaw. Unpublished material.
- Lane, P. (2010). "We did what we thought was best for our children": a nexus analysis of language shift in a Kven community. *International Journal of the Sociology of Language*, 2010(202), 63-78.
- Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health*, 13(4), 198-206.
- Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT press.
- Leonard, L. B., & Deevy, P. (2004). Lexical deficits in specific language impairment. In L. Verhoeven & H. v. Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 209). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, 51(5), 612-620.
- Łuniewska, M., Kołak, J., & Kacprzak, A. (unpublished manuscript). Polish PABiQ factors.
- Lyster, S., Horn, E., & Rygvold, A. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge: Resultater fra en utprøving av British Picture Vocabulary Scale, Second edition (BPVS II) [Vocabulary and vocabulary development of Norwegian children and youth. Results from a pilot of British Picture Vocabulary Scale, Second edition (BPVT II)]. *Spesialpedagogikk*, 74, 35-43.
- Meisel, J. M. (2004). The bilingual child. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 91-113). Malden, MA: Blackwell Publishing Ltd.
- Melby-Lervåg, M., & Lervåg, A. (2013). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Journal of Research in Reading*, 34(1), 114-135.

- Mills, M., Präg, P., Tsang, F., Begall, K., Derbyshire, J., Kohle, L., . . . Hoorens, S. (2014). *Use of childcare in the EU Member States and progress towards the Barcelona targets*. Retrieved from Brussels, Belgium: http://ec.europa.eu/justice/gender-equality/files/documents/140502_gender_equality_workforce_ssr1_en.pdf
- O'Toole, C., Gatt, D., Hickey, T. M., Miękisz, A., Haman, E., Armon-Lotem, S., . . . Kern, S. (2016). Parent report of early lexical production in bilingual children: a cross-linguistic CDI comparison. *International Journal of Bilingual Education and Bilingualism*. doi:10.1080/13670050.2016.1179258
- Office for National Statistics. (2016). *Population of the UK by Country of Birth and Nationality: 2015*. Statistical bulletin. Retrieved from <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/ukpopulationbycountryofbirthandnationality/august2016>.
- Oller, D. K., & Eilers, R. E. (2002). *Language and literacy in bilingual children* (Vol. 2). Bristol: Multilingual Matters.
- Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics*, 31(02), 227-252.
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic approaches to bilingualism*, 1(3), 213-237.
- Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research*, 59(1), 171-182.
- Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *J. of Communication Disorders*, 43(6), 474-497.
- Pearson, B. Z. (2010). We can no longer afford a monolingual norm. *Applied Psycholinguistics*, 31(02), 339-343.
- Pearson, B. Z., Fernandez, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language learning*, 43(1), 93-120.
- Peña, E. D. (2007). Lost in Translation: Methodological Considerations in Cross - Cultural Research. *Child development*, 78(4), 1255-1264.
- Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, 41(6), 1398-1411.
- Rice, M. L., & Hoffman, L. (2015). Predicting Vocabulary Growth in Children With and Without Specific Language Impairment: A Longitudinal Study From 2;6 to 21 Years of Age. *Journal of Speech, Language and Hearing Research*, 58(2), 345-359. doi:http://doi.org/10.1044/2015_JSLHR-L-14-0150

- Saltarelli, M., & Gonzo, S. (1977). Migrant languages: Linguistic change in progress. In M. de Greve & E. Rosseel (Eds.), *Problemes linguistiques des enfants de travailleurs migrants (Linguistic Problems the Children of Migrant Workers)* (pp. 167-186). Brussel, Belgium: Didier.
- Simonsen, H. G., Hansen, P., & Luniewska, M. (2012). *Crosslinguistic Lexical Tasks: Norwegian version (CLT-NO)*. A part of LITMUS Battery (COST IS0804). Department of linguistics and Scandinavian studies, University of Oslo. Unpublished material.
- Statistics Norway. (2016). Key figures for immigration and immigrants. Retrieved from <https://www.ssb.no/en/innvandring-og-innvandrerer/nokkeltall>
- Team, R. (2015). RStudio: Integrated Development for R. Boston, MA: RStudio Inc. Retrieved from <http://www.rstudio.org/>
- Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. d. Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: disentangling bilingualism from Language Impairment*. Bristol: Multilingual Matters.
- Tuller, L., Abboud, L., Ferré, S., Fleckstein, A., Prévost, P., dos Santos, C., . . . Zebib, R. (2013). Specific language impairment and bilingualism: Assembling the pieces. In C. Hamann & E. Ruigendijk (Eds.), *Language Acquisition and Development: Proceedings of GALA 2013*. Newcastle, UK: Cambridge Scholars Publishing.
- Williams, C. J., & McLeod, S. (2012). Speech-language pathologists' assessment and intervention practices with multilingual children. *International Journal of Speech-Language Pathology*, *14*(3), 292-305.

APPENDIX 1

The section on current skills from the background questionnaire, with translations to English.

<p>Czy myśli Pani/Pan, że dziecko mówi tak jak rówieśnicy, którzy znają tylko język...? <i>0 = zdecydowanie gorzej; 1 = trochę gorzej</i> <i>2 = bardzo podobnie; 3 = lepiej niż inne dzieci</i></p>	<p>Compared to other children of the same age who speak only (language), how do you think your child speaks the language? <i>0 = significantly worse; 1 = a bit worse;</i> <i>2 = very similar; 3 = better than other children</i></p>
<p>Jak Pani/Pana zdaniem dziecko wymawia słowa w danym języku w porównaniu z innym dziećmi w tym samym wieku? <i>0 = zdecydowanie gorzej; 1 = trochę gorzej</i> <i>2 = bardzo podobnie; 3 = lepiej niż inne dzieci</i></p>	<p>Compared to other children the same age, how do you think your child pronounces words in the given language? <i>0 = significantly worse; 1 = a bit worse; 2 = very similar; 3 = better than other children</i></p>
<p>Ile Pani/Pana dziecko zna słów w danym języku w porównaniu z innymi dziećmi w tym samym wieku? <i>0 = zdecydowanie mniej; 1 = trochę mniej</i> <i>2 = tyle samo; 3 = więcej niż inne dzieci</i></p>	<p>Compared to other children the same age, how many words does your child know in the given language? <i>0 = significantly fewer; 1 = a bit fewer; 2 = as many as them; 3 = more than other children</i></p>
<p>Czy Pani/Pana rodzinie i przyjaciołom łatwo prowadzić rozmowę z dzieckiem w danym języku? Czy zawsze? <i>0 = bardzo trudno; 1 = czasem są z tym problemy</i> <i>2 = zazwyczaj łatwo/łatwo</i> <i>3 = bardzo łatwo/nie ma problemów</i></p>	<p>Is it easy for your family and friends to have a conversation with your child in the given language? Always? <i>0 = very difficult; 1 = sometimes we experience difficulties; 2 = generally easy/easy; 3 = very easy/no difficulties</i></p>
<p>Czy w porównaniu z innymi dziećmi w tym samym wieku Pani/Pana dziecko radzi sobie z tworzeniem poprawnych zdań? <i>0 = zdecydowanie gorzej; 1 = trochę gorzej</i> <i>2 = bardzo podobnie; 3 = lepiej niż inne dzieci</i></p>	<p>Compared to other children the same age, do you think your child has difficulties making correct sentences? <i>0 = significantly worse; 1 = a bit worse; 2 = very similar; 3 = better than other children</i></p>
<p>Czy jest Pani/Pan zawsze zadowolona/zadowolony z tego, jak dziecko rozumie zdania, które wypowiadają do niego inne osoby w danym języku? <i>0 = zupełnie niezadowolona/niezadowolony</i> <i>1 = nie całkiem zadowolona/zadowolony</i> <i>2 = raczej zadowolona/zadowolony</i> <i>3 = całkowicie zadowolona/zadowolony</i></p>	<p>Are you always satisfied with your child's ability to understand sentences spoken to him/her by other speakers of this language? <i>0 = not at all satisfied; 1 = not very satisfied; 2 = pretty satisfied/generally satisfied; 3 = very/totally satisfied</i></p>
<p>Czy jest Pani/Pan zadowolona/zadowolony z umiejętności mówienia dziecka w danym języku? <i>0 = zupełnie niezadowolona/zadowolony</i> <i>1 = nie całkiem zadowolona/zadowolony</i> <i>2 = raczej zadowolona/zadowolony</i> <i>3 = całkowicie zadowolona/zadowolony</i></p>	<p>Are you satisfied with your child's ability to speak the given language? <i>0 = not at all satisfied; 1 = not very satisfied; 2 = pretty satisfied/generally satisfied; 3 = very/totally satisfied</i></p>
<p>Czy dziecko denerwuje się, że nie umie się porozumieć w danym języku? <i>0 = bardzo /prawie zawsze; 1 = często</i> <i>2 = czasami; 3 = prawie nigdy</i></p>	<p>Does your child feel frustrated that he/she can't communicate in the given language? <i>0 = very/almost always frustrated; 1 = often frustrated; 2 = sometimes; 3 = (almost) never frustrated</i></p>