Probability of Heritage Language Use at a Supportive Early Childhood Setting in Australia

Paola Escudero1,2*, Criss Jones Díaz3, John Hajek2,4, Gillian Wigglesworth2,4 and Eline Adrianne Smit1,2

1 The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, NSW, Australia, 2 ARC Centre of Excellence for the Dynamics of Language, The Australian National University ACT, Canberra, ACT, Australia, 3 School of Education, Western Sydney University, Sydney, NSW, Australia, 4 School of Languages and Linguistics, University of Melbourne, Melbourne, VIC, Australia

Despite well-established research that documents the intellectual, linguistic, sociocultural and familiar benefits of early childhood bilingualism, Australia’s provision of heritage language (HL) support in early childhood (EC) settings is fairly minimal, resulting in little to no access to the HL outside of the home. We report on language data from a long day care preschool that has an open bilingual policy, where two languages (HL and English) are used in naturalistic interactions between children and educators. While the quantity of language input and output is known to impact on language proficiency, there are no prior studies which focus on establishing the quantitative nature of naturalistic language production in a bilingual preschool. Our goal was to document the relative language input and output of HL and English and to examine whether there are differences across age groups in the EC setting, and during different activity types. We followed a quantitative approach in data analysis, with child and educator observations over a period of 8 weeks and an analysis of targeted videos amounting to close to an hour of recordings per child. We used Bayesian modeling to test the probability of HL use in the different age groups and per activity type. Overall, HL input was higher for toddlers than preschoolers and toddlers received more HL input than English, while preschoolers received comparable input in both languages. The higher probability of HL input in toddlers was particularly evident during story time and playing activities. Our results indicate a high level of HL use in this EC setting, suggesting success in HL maintenance and promotion of early bilingualism. Further research should explore the children’s relative language output in relation to their input, individual differences, as well as extending the current methodology to other similar settings in Australia.

Keywords: bilingualism, bilingual education, early childhood, heritage languages, relative language use, bayesian modeling

INTRODUCTION

Like many English-speaking countries, Australia has an estimated 6.8 million overseas-born inhabitants representing 35% of Australia’s population with over 21% speaking a language other than English at home, with even greater numbers doing so in large urban centers (ABS, 2016). Much of Australia's diversity lies in the cultural and linguistic repertoires of children growing up in bi/multilingual families and communities (Jones Díaz, 2018). Ideally, the Australian reality
should lead to the active practice of government policy which favors multilingualism, but this has hardly been the case (Nicholas, 2014; Eisenchlas and Schalley, 2019). This situation is, however, slowly changing for school-age children, with the increasing provision of community (or heritage) language teaching in schools and in community-based programs (Cardona et al., 2008). Unfortunately, such programs have not traditionally been accessible to most preschool children.

Despite well-established research that documents the intellectual, linguistic, sociocultural, and familial benefits of early childhood bilingualism (Blom et al., 2014; Nicolay and Poncelet, 2015; Chang, 2016), Australia’s provision of heritage language (HL) support in early childhood (EC) settings, such as preschool, remains limited, as well as little researched with children typically learning or being exposed to English at the expense of their HL. This can result in subtractive bilingualism (cf. Fillmore, 1991; Cummins, 2014; Verdon et al., 2014).

While the quantity of language input and output is known to impact on language proficiency (Hoff et al., 2014; Cha and Goldenberg, 2015; see also below), there are no prior studies which focus on establishing the quantitative nature of naturalistic language production in bilingual EC settings. In the present study, we analyse the linguistic input and output of children attending a long day care preschool that has an open bilingual policy, where two languages (HL and English) are used in naturalistic interactions between children and educators who are native speakers of the HL and fluent also in English.¹ This quantification can elucidate to what extent the language input received in a bilingual EC setting may, for instance, predict children’s output (cf. Soderstrom and Wittebolle, 2013). It may also allow researchers to estimate (a) children’s proficiency in the HL, given that input in both languages is crucial for the development of bilingualism (Hurtado et al., 2014; Place and Hoff, 2016; Unsworth, 2016; Sun et al., 2020), and (b) how successful a bilingual setting is in helping to support HL maintenance. Ultimately, although beyond the scope of this study, such information will also allow for more fine-grained analyses to be undertaken to determine which activities and specific pedagogical practices in the EC setting have the greatest impact on the children’s bilingual development.

**Benefits of Child Bi/Multilingualism**

The positive intellectual and linguistic advantages associated with child bilingualism are now widely recognized with studies demonstrating greater mental flexibility, an enhanced ability to think abstractly and to separate word referents as well as sophisticated concept formation. Two of the most significant contributions to our understanding of the linguistic benefits of bilingualism, particularly for educators, are Cummins’ (1991, 1993) “threshold hypothesis” and “interdependence hypothesis.” The “threshold hypothesis” proposes that in order for bilingual children to benefit from the cognitive and linguistic advantages of bilingualism, they must attain adequate levels of proficiency in both of their languages. The “interdependence hypothesis” proposes that there is important linguistic transfer between the two languages, and additional language learning is partly dependent upon conceptual development and proficiency already achieved in the first language. This means that in order for children to achieve high levels of majority language proficiency, both languages must be sufficiently maintained. More recent studies have advanced these findings detailing the different cognitive and metalinguistic advantages, and creative thinking resulting from child bilingualism (see for example, Cummins, 2014; Bialystok, 2015; Nicolay and Poncelet, 2015).

Studies from various perspectives have documented numerous intellectual, linguistic, cultural, social, familial and economic benefits of early childhood bilingualism, particularly in relation to children’s conceptual development (Cummins, 1991, 1993), languages and literacy learning (Hornberger and Link, 2012; Papastefanou et al., 2019), early biliteracy (Reyes and Azura, 2008; Kenner and Gregory, 2013); cultural and bi/multilingual identity (Romaine, 2011; Jones Diaz, 2016), equity and language rights (Skutnabb-Kangas, 1988; Heller, 2007; Romaine, 2013; Jones Diaz, 2018), family cohesion (Fillmore, 1991, 2000; Kouritzin, 1999; Cummins, 2009), and family language policy (Spolsky, 2004; Schwartz, 2010). Recent studies have even shown an advantage in verbal fluency in the majority language for bilingual children who gain equal levels of vocabulary in both languages (Pino Escobar et al., 2018), and in general academic skills in the majority language for bilingual children with high HL use (Papastefanou et al., 2019).

There is also a close relationship between bi/multilingualism and children’s socio-emotional trajectories with research indicating that bi/multilingual children who continue to speak their HL with family members exhibit positive socio-emotional well-being (Cummins, 2009; Genese, 2009; Han, 2010; De Houwer, 2015). For example, Han’s (2010) research investigated the connections between bilingualism and children’s development in the early years and found that most bilingual Latino/a children did as well as, if not better than, their monolingual English-speaking peers on various levels of socio-emotional well-being.

However, the role of EC education in maintaining and extending children’s HL, and the outcomes of these studies, tend to be unknown to educators of young bi/multilingual children, often resulting in deficit views toward bi/multilingual children and their families. These can lead to silences around families’ abilities to raise concerns regarding their children’s bi/multilingual development (Cline and Necochea, 2001; Pacini-Ketchabaw and Armstrong de Almeida, 2006; Robinson and Jones Diaz, 2016). When bi/multilingual children enter English-only EC settings, they may initially undergo processes of adjustment and adaptation in learning English, during which time they may become aware of the legitimacy afforded to English, and the power relations that exist between English and their own HLs. This process of rapid change, defined as submersion or “sink or swim” pedagogical approaches (Ball, 2010), may influence their attitude toward the use of their HL with family and community members, often resulting in children learning English at the expense of their HL, otherwise known as subtractive bilingualism (Fillmore, 1991, 2000).

¹To maintain full anonymity of the site in question, we refer only to Heritage Language (HL) without identifying the language in question.
Growing Up Bilingual in the Early Years of Life

The majority of the world’s population is bi/multilingual (Romaine, 2013) and this appears to be a normal, everyday expression of cultural, linguistic, and social practice (Jones Díaz, 2018). Australia is one of many multicultural/multilingual countries in a world where cultural and linguistic diversity is increasing. The 2016 census reports that Australians come from 190 different countries and 300 different ancestries, with more than 300 languages spoken, including Indigenous languages (ABS, 2016). As Australia becomes increasingly more diverse both in the number of languages as well as bi/multilingual speakers—due mainly to ongoing large-scale migration—more children than ever are growing up exposed to and potentially learning a language other than English at home.

The home contexts in which children are raised can range from the common “one parent one language model” (Döpke, 1992) where the parents speak different first languages from each other (one of which may be English), to children growing up in multilingual settings where multiple languages are spoken on a daily basis and where children hear two or three languages in their family or community from birth, often with frequent code switching. Thus, the range of circumstances in which children are raised is extremely variable, and many factors impact on eventual attainment, or not, of two (or more) languages.

One factor which impacts on children’s ability to acquire two languages simultaneously is the extent to which they receive input in both languages. This has been reported by numerous studies to date. For example, different aspects of input quantity have been shown to have an effect on bilingual development such as the relative proportion of input between languages (Hurtado et al., 2014), and the interaction with a greater range of (native) speakers (Place and Hoff, 2011, 2016; Gollan et al., 2015; Unsworth, 2016). Other researchers such as De Houwer (2011) emphasize the absolute frequency or cumulative amount of input as critical.

The question of adequate HL input is especially important where the perceived status of the languages in question varies substantially. For example, in the Australian context, the national language, English, is typically perceived as the most important language for children to learn since most of their schooling, and subsequently working life, will be conducted in it (cf. Hajek and Slaughter, 2014, and also Sun et al., 2020 for a similar situation in Singapore). However, children may have grown up in families that speak another language and because bi/multilingualism has so many positive benefits (see above), we argue that it is important to maintain both English and their other language(s). As previously noted, when children begin to attend preschool or school this is often the beginning of their close engagement with the language of the wider community, i.e., English in Australia, with which they may have previously had variable or only limited contact. Preschool or school attendance may then also be the beginning of the decrease in bi/multilingual children’s use of their other language(s) and the start of their subtractive bilingualism and their shift to the majority language (English in Australia).

Bilingual Preschools

As Schwartz and Palviainen (2016, p. 603) argue, “Early childhood is a critical period in a child’s intensive social, emotional, linguistic and cognitive development, and preschool serves as the first transitional step from home to the wider social environment and socialization of the child.” It is important, therefore, to examine the extent to which children’s languages can be supported in this environment. With a focus on children in the USA attending English-speaking kindergartens, Cha and Goldenberg (2015) explored the relationship of home input in Spanish and English (via parental report) to determine the children’s level of proficiency in both languages. Whilst cautioning against simplistic conclusions about their results, they found an association between high amounts of Spanish at home and additive bilingualism, but subtractive bilingualism where there were high amounts of English at home. In Australia there are increasing numbers of children attending preschools who come with varied language backgrounds and different levels of English proficiency. Unfortunately, we know little about the impact that maintaining these children’s HLs in these more institutional settings may have, particularly in Australia. In that respect, results provided by Fillmore (1991) show that in the US early childhood bilingual programs reduce the shift to family use of English in the home when compared to English-only early childhood programs.

Hammer et al. (2011) argue that there are four main research priorities which need to be addressed to understand children’s language development in these contexts, namely (a) studies which identify the patterns of the development of the languages; (b) longitudinal studies which document bilingual children’s language development; (c) environmental factors which may impact on language and literacy development; and (d) studies which explore children’s family backgrounds which may allow the development of maximally successful programs. The present study aligns most specifically with (c), i.e., the impact of environmental factors, specifically input.

In an experimental study which explored the outcomes of a Spanish program of language instruction on the language development of bilingual children attending an English-only preschool in the US, Restrepo et al. (2010) provided 15 Spanish L1 children who spoke only Spanish at home with either 30 min of Spanish instruction a week, for a period of 16 weeks, or with English only instruction. This small amount of instruction in Spanish enabled increased Spanish language development for the 15 students who showed increased length and complexity in their Spanish, an outcome which was not reflected in the English only group.

The present study presents a quantitative description of the relative language use of English and an HL in an EC setting in Australia to document how the two languages were used in this setting, which overtly promotes a balanced use of two languages. To our knowledge a quantitative analysis of the relative use of HL and the majority language has yet to be conducted, although recent qualitative studies of two bilingual preschools in Australia, namely a bilingual German-English in New South Wales (Benz, 2017) and a Samoan-English day care
in Queensland (Taylor-Leech, 2019), point to the prevalence and dominance of English (the majority language). From these studies, it seems that the culture and values of the HL are enhanced but not necessarily through the use of the HL in the EC settings.

In the present study, the EC setting we explore promotes the use of both the HL and English equally in statements as required by its policy guidelines and is therefore an interesting target to determine to which extent the majority language is dominant across this bilingual setting, as reported by Benz (2017), and Taylor-Leech (2019). We adopted a quantitative approach to the analysis of participant observations over a period of 8 weeks with a selection of targeted video recordings. Our analysis of daily activities was inspired by recent studies of language use in monolingual (English) preschools. Soderstrom and Wittebolle (2013) conducted a study comparing the amount of children’s vocalizations at home and in day care centers across a range of activities that tend to be present at both settings. They found that children produced the most output during play time and activities that afforded them the highest level of freedom to express themselves. In our analysis, we included similar activities to those reported in Soderstrom and Wittebolle (2013), namely play, group, meal, and story times. We chose to use targeted recordings of children and educator interactions instead of an automatic recording system such as LENA (Gilkerson and Richards, 2008) due to the bilingual nature of the study and the inability to tag the specific language in which an utterance is produced using LENA’s software.

**METHODS**

The study was conducted at an EC setting in metropolitan Sydney, in the Australian state of New South Wales. The setting was a bilingual long day care and preschool environment for children from 6 months to 5 years of age, whose program is based on a multicultural education philosophy where the main languages used are English and the HL, with the majority of educators proficient in the HL. The setting has an HL pedagogical policy with the expressed aim to provide opportunities to learn the HL. Articulated as procedure is the greater use of the HL in morning and afternoon group time, and for children to be spoken to in both HL and/or English. In relation to interactions in the HL, mealtimes are highlighted as an example of how the HL might be used in different moments throughout the day.

In this preliminary study, we focus on and document the use of children’s HL, English or a mixture of both languages as input and output in this EC setting. We report on the relative language use recorded in video footage that was taken across typical daily activities at the setting and provide a statistical analysis that models the probability of language use across the age groups and between the different activities. Ethics approval was granted for this research by the Western Sydney University Human Research Ethics Committee (Ethics No. H12055).

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Background information for the children in the present study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td>Code</td>
</tr>
<tr>
<td>Toddler T1</td>
<td>T1</td>
</tr>
<tr>
<td>Toddler T2</td>
<td>T2</td>
</tr>
<tr>
<td>Preschool P1</td>
<td>P1</td>
</tr>
<tr>
<td>Preschool P2</td>
<td>P2</td>
</tr>
</tbody>
</table>

**Participants and Data Collection**

Two toddlers and two preschoolers were each observed for half a day a week for 8-weeks. Children’s ages and languages spoken at home (including the deidentified HL) are given in Table 1. With respect to the languages spoken at home, as can be seen in the table, children had at least one parent whose first language (L1) was the HL and in most cases the other parent’s L1 was English. Interviews with parents indicated that the HL was used at home to different degrees. Since the focus of the present study is on language use at the EC setting and not at home, we do not include any further information from the interviews on potential measures of qualitative home language use.

A bilingual HL-English research assistant (RA) visited the setting mostly in the mornings from 9.30 a.m. to 12.30 p.m. during the 8 weeks. As observed by the RA, in the afternoons many of the children were tired and some took long naps, sleeping almost until pick up time. Therefore, it was considered more appropriate to visit the setting in the mornings as most of the activities and learning experiences took place between morning tea and nap time.

Video footage included various activities conducted at different times throughout the day which took place within the setting’s daily schedule. The schedule below indicate when and where activities and routines took place including in the afternoons although these were not recorded.

**Daily Schedule**

**Mornings**
- 8.00–9.30 Arrival and drop off, outdoor play (indoors when raining)
- 9.30–10.30 Morning tea
- 10.30–11.30 Story time/indoor or outdoor play/group activity
- 11.30 Lunch

**Afternoons**
- 1.00–2:30 Nap time/or quiet reading (average: 1 h for preschoolers; 1.5 h (toddlers))
- 2:30–4:00 Outdoor play (indoors when raining)/group activity/story time
- 4.00–3.40 Afternoon tea
- 4:30–6:30 Departure and pick up

The sample video data included four morning routines: mealtimes, story time, indoor and outdoor play (e.g., free play with toys, and in the sandpit) and group activities (e.g., singing songs, making playdough, and yoga).
As this study is exploratory, and there were differences between children and adults. Our approach was to select videos that amounted to similar durations per activity and per child, as shown in Table 2. In addition to the 48 videos selected for the activities—meals, playing, and group activity—we included 13 videos from story time which involved the entire story time, including giving directions for gatherings, singing, and speeches, that took place during the story time observations. Seven of these videos were stories read to the toddlers and six were stories read to the preschoolers. The total duration of all selected videos amounted to approximately 1 h of video footage per focus child. There was a total of 138 short videos (with various durations) that were taken throughout eight observation visits. Sixty-one videos (44% of the total) were selected for analysis.

As shown in Table 2, the amount of video recordings per minute for each child were similar except for mealtimes which were much shorter for the preschool children than for the toddler group because less interaction took place during the preschoolers' mealtimes.

### Data Analysis

In order to calculate the relative language use at the setting, the bilingual RA examined each video clip and noted in a spreadsheet the length of time (in seconds) that an adult or a child spoke in either language (or a mixture of the two languages) during each video. First, we coded the participant ID, the start and end of each utterance produced by either an adult or a focus child within the specific video, and the language in which the utterance was produced as well as silences. Adult speech to a focus child, or speech by another focus child in the study, were coded as input. Speech by the focus child was coded as output. For example, video 27 had preschooler 1 as the focus child. The activity in this video was "play" which lasted for 1 min and 11 s. In this video, an adult spoke in English once for 9 s, and then in the HL five times for a total of 16 s. Preschooler 1 spoke in English for 19 s and in the HL for 2 s. This amounts to 16 s of HL input, 9 s of English input, 2 s of HL output, and 19 s of English output.

### Statistical Analysis

We used a Bayesian regression model to calculate the probability that speech output is in either of the two languages (English or HL). The Bayesian analysis was conducted using the brms package in R (R Core Team, 2014; Bürkner, 2017, 2018). A similar approach has been used previously in Smit et al. (2019) who also provides a short overview of Bayesian regression (see also van de Schoot and Depaoli, 2014; van de Schoot et al., 2014). Bayesian statistics have a number of benefits compared to conventional frequentist methods, such as linear mixed models or analysis of variance (van de Schoot and Depaoli, 2014), that are particularly suited for naturalistic child data. Data collection did not take place in a controlled laboratory, but at a noisy bilingual preschool, where variables were neither controlled nor balanced, and thus they did not meet normality assumptions so that using conventional frequentist methods would be difficult and could comprise the results. Importantly, the sample size for the children's speech was quite small compared to the data collected for educators. Unlike linear mixed models, Bayesian models can be computed despite the issues mentioned above (van de Schoot and Depaoli, 2014).

Another important benefit of Bayesian statistics is that they allow for the incorporation of background knowledge obtained from comparable prior research. In this way, cumulative meta-analytic knowledge can be built, which cannot be done with the frequentist approach of NHST (null-hypothesis significant testing) (König and van de Schoot, 2018). This background knowledge is reflected in the models by the use of priors. Before analyzing the data, a prior distribution can be chosen which reflects the amount of (un)certainty in the model's parameters (Smit et al., 2019). As this study is exploratory, and there were no prior assumptions about the behavior of the data's parameters, we used the so called weakly informative priors (Gelman et al., 2014, 2015). Crucially, the results presented here can be reflected in the priors that can be chosen in similar future research. For the current analysis, we used weakly informative priors with a Student's t-distribution with 3 degrees of freedom, a mean of 0, and a scale of 2.5.

In Bayesian statistics, uncertainty is defined by the probability distribution of the population parameter (van de Schoot et al., 2014). Unknown parameters (such as the population mean) are thus not considered to be one fixed value (as is the case in frequentist models), but an interval describing the probability that the parameter's true value of a predictor variable (i.e., an independent variable) is within this interval, known as a credibility interval. For this study, we calculated 95% credibility intervals, thus showing the 95% certainty that a parameter's true value falls within that region.
We can test whether one parameter’s interval differs from another using a hypothesis-testing method, resulting in evidence ratios that quantify the likelihood of a given hypothesis with respect to the alternative. Following Jeffreys (1961, as cited in Kruschke, 2018), evidence ratios of 3–10 represent “moderate” evidence for a given hypothesis, 10–30 represent “strong” evidence, while ratios above 30 represent “very strong” evidence for a given hypothesis. In contrast, evidence ratios of 1/3–1/10 are “moderate” evidence against a hypothesis; 1/10–1/30 “strong” evidence against and evidence ratios beyond 1/30 are “very strong” evidence against a hypothesis. For reference, evidence ratios of 19 are roughly similar to a p-value of 0.05 in NHST (Milne and Herff, in press).

We used language output (HL, English, or mixed) as the categorical dependent variable, weighted by the number of seconds a language was used during the interactions recorded in the selected videos. Predictor variables were Group (toddler or preschool), Activity (story time, meals, playing, or group activity), and Speaker (Educator 1–9 and Child 1–4). We fitted a model to predict the language output per Group, Activity, and Speaker, with interactions between all predictors. Predictors were turned into factors and referenced to English (for Language), Toddlers (for Group), and Meals (for Activity). Models converged as evidenced by Rhat values below 1.1, meaning that sufficient iterations were run and that the chosen priors were sufficiently strong (Bürkner, 2017).

**RESULTS**

Below we detail our quantitative results. We first review the results from the descriptive analysis of the raw data, followed by the results of the Bayesian regression model which calculated the probability of speech output being in each of the two languages (English or HL). As this study is in the first instance descriptive and exploratory in nature, we did not formulate any hypotheses about language use a priori, but based on the descriptive results, we were then able to formulate a number of hypotheses that were then tested against the data. We averaged results per age group because due to the nature of our naturalistic data collection, children from the same age group appeared in many videos together and educators commonly addressed all children in an age group at the same time. As well, since the aim of the study was to examine the relative language use in the setting, we focused on age groups rather than individual children.

**Descriptive Results**

Table 3 summarizes the results for educators’ language input to children and output by children across the different languages, English, HL, or mixed utterances where both languages were used in a single utterance.

<table>
<thead>
<tr>
<th>Language</th>
<th>Age group</th>
<th>Educator input to children</th>
<th>Output by children</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Toddlers</td>
<td>23.80</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Preschoolers</td>
<td>23.70</td>
<td>5.48</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>47.50</td>
<td>6.40</td>
</tr>
<tr>
<td>HL</td>
<td>Toddlers</td>
<td>39.62</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td>Preschoolers</td>
<td>21.43</td>
<td>4.00</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>61.08</td>
<td>6.33</td>
</tr>
<tr>
<td>Mixed</td>
<td>Toddlers</td>
<td>0.57</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Preschoolers</td>
<td>1.88</td>
<td>0.13</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>2.45</td>
<td>0.25</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>111.03</td>
<td>12.98</td>
</tr>
</tbody>
</table>

**Table 4A** | Input to toddler group in minutes.

<table>
<thead>
<tr>
<th>Toddler input</th>
<th>English</th>
<th>HL</th>
<th>Mixed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story time</td>
<td>11.63</td>
<td>22.72</td>
<td>0.00</td>
<td>34.35</td>
</tr>
<tr>
<td>Meals</td>
<td>2.60</td>
<td>4.52</td>
<td>0.18</td>
<td>7.30</td>
</tr>
<tr>
<td>Playing</td>
<td>1.22</td>
<td>5.75</td>
<td>0.00</td>
<td>6.97</td>
</tr>
<tr>
<td>Group activity</td>
<td>8.35</td>
<td>6.63</td>
<td>1.35</td>
<td>16.33</td>
</tr>
<tr>
<td>Total</td>
<td>23.80</td>
<td>39.62</td>
<td>1.53</td>
<td>64.95</td>
</tr>
</tbody>
</table>

**Table 4B** | Input to preschool group in minutes.

<table>
<thead>
<tr>
<th>Preschool input</th>
<th>English</th>
<th>HL</th>
<th>Mixed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story time</td>
<td>16.97</td>
<td>12.00</td>
<td>0.00</td>
<td>28.97</td>
</tr>
<tr>
<td>Meals</td>
<td>1.40</td>
<td>0.78</td>
<td>0.22</td>
<td>2.40</td>
</tr>
<tr>
<td>Playing</td>
<td>1.02</td>
<td>1.22</td>
<td>0.32</td>
<td>2.55</td>
</tr>
<tr>
<td>Group activity</td>
<td>4.32</td>
<td>7.43</td>
<td>1.35</td>
<td>13.10</td>
</tr>
<tr>
<td>Total</td>
<td>23.70</td>
<td>21.43</td>
<td>1.88</td>
<td>47.02</td>
</tr>
</tbody>
</table>

**Table 4C** | Output from toddler children in minutes.

<table>
<thead>
<tr>
<th>Toddler output</th>
<th>English</th>
<th>HL</th>
<th>Mixed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story time</td>
<td>0.00</td>
<td>0.78</td>
<td>0.00</td>
<td>0.78</td>
</tr>
<tr>
<td>Meals</td>
<td>0.45</td>
<td>0.23</td>
<td>0.05</td>
<td>0.73</td>
</tr>
<tr>
<td>Playing</td>
<td>0.42</td>
<td>0.55</td>
<td>0.03</td>
<td>1.00</td>
</tr>
<tr>
<td>Group activity</td>
<td>0.05</td>
<td>0.77</td>
<td>0.03</td>
<td>0.85</td>
</tr>
<tr>
<td>Total</td>
<td>0.92</td>
<td>2.33</td>
<td>0.12</td>
<td>3.37</td>
</tr>
</tbody>
</table>

**Table 4D** | Output from preschool children in minutes.

<table>
<thead>
<tr>
<th>Preschool output</th>
<th>English</th>
<th>HL</th>
<th>Mixed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story time</td>
<td>0.00</td>
<td>0.78</td>
<td>0.00</td>
<td>0.78</td>
</tr>
<tr>
<td>Meals</td>
<td>0.45</td>
<td>0.23</td>
<td>0.05</td>
<td>0.73</td>
</tr>
<tr>
<td>Playing</td>
<td>0.42</td>
<td>0.55</td>
<td>0.03</td>
<td>1.00</td>
</tr>
<tr>
<td>Group activity</td>
<td>0.05</td>
<td>0.77</td>
<td>0.03</td>
<td>0.85</td>
</tr>
<tr>
<td>Total</td>
<td>0.92</td>
<td>2.33</td>
<td>0.12</td>
<td>3.37</td>
</tr>
</tbody>
</table>

Tables 4A–D summarize the time (in minutes) that each language was used in each age group, by both educators and children, by activity.

The input to both groups varies to some degree, particularly in story time where toddlers receive much more HL input than the preschoolers, although both receive a considerable amount of English. The toddler group receives more input in general during meals and noticeably more HL input while playing than
the preschool group. Conversely, the preschool group receives very limited input during playtime and meals in either language.

Tables 4C,D indicate that there is very limited language use by the children, but that playing elicits a greater amount of language (in English) among the preschool children, consistent with results found by Soderstrom and Wittebolle (2013).

Results From the Statistical Modeling

From the patterns found in the descriptive results reported above, we formulated the following eight specific hypotheses to test the evidence that supports that the *probability* of educators’ language use in a specific age group or activity is different from the probability of their language use in another group or activity. We focus on educators’ output in HL and English only, as children’s output and mixed language speech were minimal (see Tables 4A,B).

\[ H1. \text{ Educators’ HL use in the toddler group is in general higher than in the preschool group} \]
\[ H2. \text{ Educators’ HL use is higher than English in the toddler group} \]
\[ H3. \text{ Educators’ HL use is lower than English in the preschool group} \]
\[ H4. \text{ Educators’ HL use during story time is higher for toddlers than for preschoolers} \]
\[ H5. \text{ Educator’s HL use during story time is higher for toddlers than for preschoolers} \]
\[ H6. \text{ Educator’s HL use during playing is higher than English for toddlers} \]
\[ H7. \text{ Educator’s HL use during playing is lower than English for preschoolers} \]
\[ H8. \text{ Educator’s HL use is higher than English during group activities for preschoolers} \]

The results of our hypothesis testing, including the strength of the evidence for each hypothesis, are outlined in Tables 5, 6 below. In Table 5, we can see that there is very strong evidence to show that the probability of Educators’ HL use for toddlers is indeed higher than for preschoolers (H1). When we look at the groups separately, educator input to toddlers is more likely in the HL than in English (H2). For the preschool group, there is no evidence that HL input is lower than English input (H3).

Table 6 reports on those activities that, from the descriptive results, appeared to show differential language use, namely during story time and playing for the toddlers, and story time and group activity for the preschoolers. There is very strong evidence of a higher probability of HL during story time in the toddler group compared to the preschool group (H4). For toddlers, we find very strong and strong evidence of a higher probability of HL use by educators both for story time and playing, respectively (H5 and H6). For the preschool group, there is no evidence to support a higher prevalence of HL or English for story time and group activity (H7 and H8) suggesting that the difference in educator use of language between HL and English for those activities is not large enough to be supported with statistical evidence.

Overall, the results support the descriptive statistics and indicate that there is a higher use of HL in the toddler group compared to the preschool group, in particular for story time and playing. There is not enough evidence to conclude that either HL or English is more prevalent in preschoolers.

**DISCUSSION**

This study examined the naturalistic use of HL and English language in four different activities, in two different age groups, in a bilingual preschool in Australia. The results clearly show that in this EC setting children receive substantial amounts of input (i.e., educator output) in both languages—at least relative to the output generated by the children themselves (see Table 2). Previous discussion of earlier research points to quantity of language input as a predictor of language proficiency, and we would suggest the approach taken in this EC setting supports children’s use and learning of HL and English. The fact that the children also have the opportunity to interact with a range of adult speakers in the setting, in addition to input from family members, is also likely to benefit their language proficiency, in line with previously reported findings by Gollan et al. (2015), Place and Hoff (2011, 2016) and others.

While the overall HL input provided by the educators is found to be greater than that for English, particularly in the toddler group, it was not the case that input quantity was consistent across activities or groups. It is clear from our data that there is some variation. As predicted, input in either language is greatest during story time and group activity which are highly structured teacher-led activities. The relative amount and proportion of HL

---

2 The observed results might be misrepresented due to the presence of an educator who is a native English speaker in the preschool group. We repeated the same analyses excluding the language output of this educator and found that the results did not change.
and English also varies according to activity and group, e.g., there is greater absolute and relative use of HL than English in story time with toddlers than with preschoolers, while during mealtimes input in English and HL is less but more evenly divided. Clearly, a balance of input and output is necessary in order to maximize children’s linguistic capabilities in terms of asking questions, giving information, expressing likes and dislikes, so that interactions are meaningful and communicative. However, our recordings may not have captured the full extent of children’s output at the EC setting, which may for instance require recording more peer-based play time or elicitation activities, especially for the preschoolers.

Our statistical modeling of the educators’ relative language use confirmed that the probability of HL use was higher with toddlers than with preschoolers and that the HL was used more than English in toddlers but not in preschoolers (where both English and HL were used to comparable extents). These findings confirm an increase in the relative use of English with age but also confirm that the EC setting is successful in providing ample opportunities for the children to receive equal or even more input in the HL, supporting HL maintenance.

There were insufficient data to model the children’s output, but based on the trends found in the raw data, we suggest this shift may be an indication of a gradual transition to English, as the children in the older preschool group begin to move toward the expected use of English in formal schooling. The fact that output in English is by far the greatest during preschoolers’ play may also be an effect of typical inter-peer socialization in Australia (as in other English-speaking countries) which generally favors the rapid adoption and use of English amongst young children as they become older, especially if they have contact with siblings at home and other children who are already at school (cf. Michael-Luna, 2013). Interestingly, it is during play that the preschoolers also show the greatest HL output, albeit alongside even greater use of English. The relatively high levels of HL output observed in this specific context may suggest that the EC setting has still been able to engage children successfully in terms of HL output. However, it is important to note that more data would be needed in order to statistically test these assumptions.

Despite the group focus of the current analysis, further investigation would be useful to understand the home setting of individual children (e.g., the absence or presence of older siblings) or whether there are specific strategies used by the staff which support HL output. Li’s (2012) study into the role of imaginative play in bilingual HL development confirms the importance of communicative and interactive linguistic strategies used by adults. While her study focussed on Mandarin-speaking parents’ role in extending their children’s Mandarin in play-based contexts, the findings highlight play as an affective pedagogical tool which affords strategies such as questioning, negotiating, and extending on play dialogues, between children and adults to enhance the imagined situation to broaden the play experience. Through this process, bilingual children’s language use in both languages is extended.

With respect to children’s overall output, we find relatively small amounts and proportions in both languages compared to the input the children received from the educators—regardless of the group setting. The high ratio (8.55 to 1) overall of adult input to children’s output previously noted suggests that substantial input from adults may be needed to generate children’s output in any language. However, some caution is needed here: structured teacher-led activities such as story time and group activities may not necessarily require active language production from the children because of the activity underway, e.g., listening to a story, or to the instructions for a game. Nonetheless, considerable effort must be given to providing children with input since this also provides rich language through modeling, scaffolding, and questioning to extend children’s vocabulary, and grammatical structures for children to use and master.

Importantly, if the amount of HL input is known to be a predictor of proficient HL output, then targeted and overt educational strategies designed to maximize children’s use of the HL may be necessary.

Bi/multilingual parents may be unsure of how far to support HL maintenance by their children. They often express anxiety about language mixing, decisions about which language to use at home, and reading books in two (or more) languages fearing a language deficiency will impact negatively on their children’s learning not only of the HL but critically of English (Michael-Luna, 2013). Our data from naturalistic speech show, however, language mixing is minimal and does not appear to be a significant feature of this EC setting.

**TABLE 6 | Results from the hypothesis tests for educator output per selected activities.**

<table>
<thead>
<tr>
<th>Language probability</th>
<th>Est.</th>
<th>Est. Error</th>
<th>95% CI</th>
<th>Evid. Ratio</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Toddler vs. preschooler</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H4 Story time</td>
<td>HL &gt; English</td>
<td>10.08</td>
<td>6.75</td>
<td>1.47–Inf.</td>
<td>37.10</td>
</tr>
<tr>
<td><strong>Toddler</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H5 Story time</td>
<td>HL &gt; English</td>
<td>8.84</td>
<td>5.50</td>
<td>3.08–Inf.</td>
<td>180.82</td>
</tr>
<tr>
<td>H6 Playing</td>
<td>HL &gt; English</td>
<td>8.59</td>
<td>4.58</td>
<td>2.74–Inf.</td>
<td>171.41</td>
</tr>
<tr>
<td><strong>Preschooler</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H7 Story time</td>
<td>HL &lt; English</td>
<td>−1.24</td>
<td>3.91</td>
<td>−Inf. to 4.32</td>
<td>1.67</td>
</tr>
<tr>
<td>H8 Group activities</td>
<td>HL &gt; English</td>
<td>2.86</td>
<td>5.93</td>
<td>−5.12 to Inf.</td>
<td>2.20</td>
</tr>
</tbody>
</table>

*Est. (estimate) = mean coefficient. Estimate error = standard deviation of the posterior distribution. 95% CI are two-sided 95% credibility intervals. Evid. Ratio (evidence ratio) = the posterior probability under the hypothesis against its alternative. Evidence = strength of the obtained evidence.*
Although individual differences in terms of language choice across activities and reasons for them are beyond the scope of this study, they deserve further investigation. Factors affecting individual differences in children may include home environment in terms of parental strategies for HL maintenance, older siblings’ use of English at home, children’s emerging identity as bi/multilingual speakers, children’s attitude toward their HL, parental attitudes, and expectations of the setting, educators’ professional development opportunities in facilitating bi/multilingual children’s linguistic repertoires, and children’s affinity with other children and educators among others (Ball, 2010; Schwartz, 2010; Jones Díaz, 2011, 2014, 2018; Benz, 2017). Although they are likely to affect children’s output, home environment differences may also influence educators’ language choice when addressing specific children. However, since they tend to address a group rather than individual children, educators’ individual choice may be better explained by their educational strategy, and the settings’ language policy and their linguistic background.

CONCLUSION

In this paper we have documented the input and output of four children attending a bilingual preschool in Australia. The documentation demonstrates that unlike in previous reports of bilingual EC settings in Australia where English is predominant (Benz, 2017; Taylor-Leech, 2019), these children receive a considerable amount of their input in both the HL and English in the various activities with which they engage during the day, and for the most part there is also output in both languages from the children. Crucially, younger children receive more HL than English input and they are more likely to hear HL spoken in the classroom than preschool children, indicating that age predicts the amount of HL provided in this EC setting. Overall, these findings suggest that this setting is successful in promoting bilingualism and HL maintenance in early childhood.

These findings highlight the need for further research into the role of EC education in maintaining and extending children’s HLs outside of the home. Further research should examine and compare multiple similar EC settings and examine the quality of input and output in each language, which would yield rich insights into the children’s bi/multilingual development, and the role of EC educators in providing opportunities for extended use of the HL and English in different programs. Such research can not only inform curriculum and policy, but provide guidance to bi/multilingual families as to home-based effective interactional and communicative strategies that support their children’s bi/multilingual development.

REFERENCES


DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Western Sydney University Human Research Ethics Committee (Ethics No. H12055). Written informed consent to participate in this study was provided by the participants’ legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

PE and CJ conceived the qualitative component of the project and developed the observational data collection protocol. PE conceived the quantitative component of the project reported and the time-span analysis of relative language use reported in this paper. PE and ES conceived the statistical analysis method. ES computed descriptive statistics and ran the statistical model from the data that was coded by an RA and checked by a research officer with input from PE, JH, and GW. PE, CJ, JH, GW, and ES wrote the paper. All authors contributed to the article and approved the submitted version.

FUNDING

The initial stage of the project including a literature review and a survey was funded by an internal grant from the MARCS Institute for Brain Behavior and Development at Western Sydney University. Project development and data collection were funded by a Seed Grant from the Centre for Educational Research from the School of Education at Western Sydney University. Data coding, analysis, and publication fees were funded by the Australian Research Council (ARC) Centre of Excellence of the Dynamics of Language (CE140100041) of which PE and GW were Chief Investigators and JH and ES are affiliates. PE’s work was supported by an ARC Future Fellowship (FT160100514).

ACKNOWLEDGMENTS

We thank Gloria Pino Escobar for her help with our ethics application, recruitment, data collection and data coding. We thank Dr. Weicong Li for his help developing the data coding scheme and his initial analysis of the coded data, and Dr. Beatriz Cardona for help with article formatting. We also thank Dr. Andrew Milne for his advice regarding the data modeling. Special thanks to the director, educators, parents, and children who took part in the study.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Western Sydney University Human Research Ethics Committee (Ethics No. H12055). Written informed consent to participate in this study was provided by the participants’ legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

PE and CJ conceived the qualitative component of the project and developed the observational data collection protocol. PE conceived the quantitative component of the project reported and the time-span analysis of relative language use reported in this paper. PE and ES conceived the statistical analysis method. ES computed descriptive statistics and ran the statistical model from the data that was coded by an RA and checked by a research officer with input from PE, JH, and GW. PE, CJ, JH, GW, and ES wrote the paper. All authors contributed to the article and approved the submitted version.

FUNDING

The initial stage of the project including a literature review and a survey was funded by an internal grant from the MARCS Institute for Brain Behavior and Development at Western Sydney University. Project development and data collection were funded by a Seed Grant from the Centre for Educational Research from the School of Education at Western Sydney University. Data coding, analysis, and publication fees were funded by the Australian Research Council (ARC) Centre of Excellence of the Dynamics of Language (CE140100041) of which PE and GW were Chief Investigators and JH and ES are affiliates. PE’s work was supported by an ARC Future Fellowship (FT160100514).

ACKNOWLEDGMENTS

We thank Gloria Pino Escobar for her help with our ethics application, recruitment, data collection and data coding. We thank Dr. Weicong Li for his help developing the data coding scheme and his initial analysis of the coded data, and Dr. Beatriz Cardona for help with article formatting. We also thank Dr. Andrew Milne for his advice regarding the data modeling. Special thanks to the director, educators, parents, and children who took part in the study.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Escudero, P; Jones Diaz, C; Hajek, J; Wigglesworth, G; Smit, EA

Title:
Probability of Heritage Language Use at a Supportive Early Childhood Setting in Australia

Date:
2020-07-16

Citation:

Persistent Link:
http://hdl.handle.net/11343/241496

File Description:
Published version

License:
CC BY