Individual Differences in the Lexical Development of French–English Bilingual Children

Annabelle David
School of Modern Languages, Newcastle University, Newcastle upon Tyne, UK

Li Wei
School of Languages, Linguistics & Culture, Birkbeck, University of London, London, UK

The large and rapidly expanding body of literature on bilingual acquisition is mostly comprised of either single-case or cross-sectional studies. While these studies have made major contributions to our understanding of bilingual children’s language development, they do not allow researchers to compare and contrast results with regard to individual differences over time. This paper aims to investigate the issue of individual differences with a longitudinal group study of 13 French–English bilingual children. The main focus is lexical development. We will examine how extralinguistic factors such as gender, parental input and birth order impact on the lexical development of the children. Using quantitative (parental checklists, questionnaires) and qualitative measures (interactions with parents), we demonstrate that language exposure and parental input are closely linked to vocabulary size, amount of language mixing and cross-linguistic synonyms. The findings call for more longitudinal group studies of bilingual acquisition in order to obtain comparable results on larger populations.

doi: 10.2167/beb478.0

Keywords: Bilingual acquisition, group study, vocabulary, French

Introduction

This paper has two closely related aims. It reports on the findings from a longitudinal group study of French–English bilingual children, focussing on lexical development over time. It also addresses the theoretical and methodological implications of such a longitudinal group study. We are particularly concerned here with the issue of individual differences in early bilingual acquisition, an issue that has hitherto received relatively little attention due to lack of longitudinal group studies.

The landscape of research on bilingual language acquisition is fast changing. Earlier focus on the one-versus-two-system(s) hypothesis (e.g. Genesee, 1989; Meisel, 1989; Volterra & Taeschner, 1978), i.e. whether bilingual infants had a ‘fused’ linguistic system, gave way to an increased interest in cross-linguistic interactions between languages in the developing bilingual repertoire (see Zhu & Li, 2005 for a review). Theoretically, separate development is possible without there being any similarity with monolingual
acquisition. However, much of the research evidence seems to suggest that bilingual children’s language development is by and large the same as that of monolingual children. In very general terms, both bilingual and monolingual children go through an initial babbling stage, followed by the one-word stage, the two-word stage, the multiword stage and the multiclause stage. At the morphosyntactic level, a number of studies have reported similarities rather than differences between bilingual and monolingual acquisition. Garcia (1983), for example, compared the use of English morpheme categories by English monolingual children and bilingual children acquiring English and Spanish simultaneously and found no systematic difference at all. Pfaff and Savas (1988) found that their four-year-old Turkish–German subject made the same errors in Turkish case marking as reported in the literature on monolingual Turkish children. Müller’s (1994) study of two French–German children suggests that their use of subject-verb agreement and finite verb placement in both languages is virtually identical to that of comparable monolingual children. De Houwer (1990) found that her Dutch–English bilingual subject, Kate, used exactly the same word orders in Dutch as monolingual Dutch-speaking children, both in terms of types and in proportional use. Furthermore, De Houwer found in Kate parallels to monolingual children for both Dutch and English in a range of structures, such as non-finite verb placement, preposed elements in affirmative sentences, clause types, sentence types, conjunctions and question inversion.

Nevertheless, one needs to be careful in the kinds of conclusions one draws from such evidence. Similarities between bilingual and monolingual acquisition do not mean that (i) the two languages a bilingual child is acquiring develop in the same way or at the same speed; (ii) the two languages a bilingual child is acquiring do not influence and interact with each other. Paradis and Genesee (1996), for example, found that the French–English bilinguals they studied (aged between 2 and 3 years old) displayed very similar developmental patterns to those of monolinguals in terms of syntactic development. Not only was their development similar but it also appeared to be within the same age range. Paradis and Genesee (1996) observed, amongst other things, the earlier use of finite verbs in French (when compared to English) and the placement of negatives after lexical verbs in French (e.g. ne regarde pas) and before lexical verbs in English (e.g. do not watch). Evidence of cross-linguistic influence has been reported by Döpke (2000), for example, in her study of German–English bilingual children in Australia. They tended to overgeneralise the -VO word order of English to German which instantiates both -VO and -OV word orders depending on the clausal structure of the utterance. Döpke suggests that children learning English and German simultaneously are prone to overgeneralise S-V-O word order in their German because the -VO order is reinforced on the surface of both the German and the English input they hear.

Vocabulary development in bilingual children has often been a topic studied from a bilingual perspective. By this we mean that the issues linked with lexical development have been mainly bilingual issues, i.e. the presence or not of translation equivalents (e.g. Deuchar & Quay, 2000; Volterra & Taeschner, 1978) and the implications for the cognitive representation of
languages in the bilingual brain. A few more recent studies have started looking at vocabulary from a more monolingual perspective, i.e. looking at similar issues in monolingual and bilingual children. Junker and Stockman (2002) found that bilingual toddlers at two were not inferior (to monolinguals) in conceptual vocabulary size and verb diversity when words in both languages were accounted for. Holowka et al. (2002) argued that their bilingual children reached the major lexical milestones within what is considered typical monolingual range. Another study, by Allman (2002), also confirmed that her Spanish–English bilinguals’ receptive and productive lexicons were statistically no different to those of monolinguals English or Spanish children.

One important methodological issue that becomes apparent even in this brief and highly selective review of literature is the fact that the vast majority of published studies of bilingual acquisition depended on single case studies. There is no question that such detailed, systematic analyses of individual cases have offered deep insights into the nature of the bilingual child’s developing language. Yet, it is very difficult to compare or to generalise the findings from single cases. Research on monolingual child language acquisition has revealed significant individual variability. For instance, gender has been reported to be an important factor in the rate of language acquisition (especially lexical development). Tamis-Lemonda et al. (1998) found a significant difference between girls and boys achieving different language milestones in their second year, girls generally being earlier than boys. Le Normand et al. (2002) found that girls’ lexical development (French) was much earlier and faster than boys under the age of two. Similarly, Bornstein et al. (1998) found significant differences between English-speaking girls and boys in various measures of vocabulary competence. Other studies using parental checklists to assess vocabulary production and comprehension also found similar differences (Eriksson, 2001; Fenson et al., 1993; Maital et al., 2000). Generally speaking, the differences become less significant as the children get older (see Kern, 2001). There are, of course, studies that did not reveal any significant differences between girls and boys. Schachter (1979), for example, found that the MLU of boys and girls was very much the same. Bates et al. (1988) claimed that differences between girls and boys in vocabulary acquisition existed but were insignificant before the age of two, and they were ‘levelled out’ if the analysis took the first three years as a whole.

Input has also been shown to be a major factor in language acquisition in general. Studies of monolingual children show that the quantity and quality of parental input differ immensely both in terms of fathers, mothers and family types and in terms of changes over time. Most studies seem to suggest that mothers and other female caregivers spend more time with young children in joint activities than fathers and other male members of the child’s immediate social network. Hart and Risley (1995) claimed that the amount of parent language varied significantly according to family types; nuclear, middle-class families seemed to provide a better language environment than large, working-class families. They also noted a rise in speech directly addressed to the child at around the time of the child’s first words. They further argued that differences in the quantity of input led to differences in ‘the richness of quality utterances’ (Hart & Risley, 1995: 127). With regard to the quality of
input and its effect on children’s language development, Shore (1995) claimed that differences in the type of input (e.g. use of formulaic utterances, speed, complexity, etc.) and what she called ‘parental responsiveness versus directiveness’ (Shore, 1995: 75) could result in significant differences in the children’s linguistic development; the more directive mothers were, the less referential the vocabulary of their children was likely to be. Bornstein et al. (1998) revealed that mothers’ own vocabulary directly influenced their children’s verbal comprehension and maternal reports of the child’s vocabulary. Similarly, Parisse and Le Normand (2000) showed that about 80% of words (in tokens) produced by a French-speaking child at 2;1 corresponded exactly to those produced by the adults he was interacting with.

Closely related to the factor of parental input is the issue of birth order. Parents are not the only source of input for a language acquiring child. Their siblings have a significant role in their overall development. There has been limited research on the effect of birth order on language acquisition. Bennett-Kastor (1988) suggested that the greater amount of time that first-borns spent in interaction with their parents, as opposed to with younger siblings, may enhance their language development. Bates et al. (1994) found ‘a small but reliable correlation’ between lexical acquisition of nouns and birth order. Later-born children had a smaller vocabulary than first-born children. Similar results were found in different languages (e.g. Kern, 2001; Maital et al., 2000).

All of these studies focused on monolingual children. But there is no reason to believe that the factors that were found to be contributing to the variations in individual children’s language development do not apply to bilingual children as well. Harding and Riley’s (1986) typology of childhood bilingualism, later expanded by Romaine (1995), who added one category, highlights the importance of sociolinguistic environment, parental input and interactional strategies in influencing the outcome of bilingual children’s language ability and behaviour. A number of studies of bilingual children allude to some of these factors. One study on lexical acquisition by Pearson et al. (1997), for example, found a strong correlation between language exposure and words known in each language by young English–Spanish bilinguals. Pearson et al. did find that there was a direct effect of time spent in each language and that ‘even at reduced levels of exposure to a language, children will still learn its vocabulary’ (Pearson et al., 1997: 55). Supporting evidence for the correlation between language exposure and lexical or morphosyntactic development in young bilinguals also comes from language attrition studies. Turian and Altenberg (1991) reported the case of the first author’s son who grew up exposed to Russian and English. When aged around 3;6, Joseph’s input in Russian decreased dramatically (from eight hours a day to just one during the week). The authors noticed that certain aspects of language the child seemed to have acquired in Russian (e.g. nominative-accusative control) were lost when the child was taped between 4;3 and 4;4. The child also acknowledged ‘forgetting’ some vocabulary items (Turian & Altenberg, 1991: 211). Yukawa (1997) analysed the lexical and grammatical development of two Japanese–Swedish bilinguals. The study showed that the children’s abilities to use certain structures and items were closely related to the changing exposure to the two languages (the children moved between Japan and Sweden). Similar
results were noted by Lanvers (1999) in a single case study on the acquisition of translation equivalents. She claimed that ‘changes of input greatly increased equivalent learning’ (Lanvers, 1999: 30).

Perhaps the best studied extralinguistic factor in bilingual language acquisition is parental discourse strategy. Lanza suggested that language choice might be affected by strategies used by parents in particular in response to instances of mixing from the child. Lanza’s findings gave rise to what Nicoladis and Genesee (1998) termed the Parental Strategy Hypothesis. However, Nicoladis and Genesee (1998) failed to reveal any kind of correlation between mixing and parental strategies in their single case study (see also Deuchar and Muntz, 2003). Lanza (2001) argued that parental discourse strategies should be used as a qualitative measure and not quantitative. This opinion is shared by another study by Juan-Garau and Perez-Vidal (2001), who did not calculate any correlations but attributed the decrease of language mixing of their child to the change of strategies used by the father.

More recently, some researchers have found evidence of birth-order effect on bilingual acquisition (e.g. Shin, 2005), with first-born children in bilingual families being more proficient in the home language than their younger siblings. As Zhu and Li (2005) noted, studies reporting on children growing up in situations where one language (the minority one) is spoken at home and the other one in the community (majority language), usually state that peers and siblings mainly adopt the majority language as their primary language of interaction, thus contrasting with that of their parents. Children born in families with older siblings usually form a separate network among themselves. That is bound to have some kind of impact on the language development of the bilingual child and should be considered as a possible and significant factor.

These studies call for more detailed, systematic analyses of individual differences in bilingual acquisition. However, such analyses could not be possible without longitudinal group studies. Single case studies by definition cannot control all the variables and make comparison across subjects difficult. The present study is one of the first group longitudinal studies of lexical development of young bilingual children. In what follows, we focus on individual differences and try to offer some tentative explanations as to what factors might have led to the differences.

**Methods**

The present study involves 13 bilingual French–English children aged between 1;0 and 3;0. All the children came from families that follow the one-parent-one-language (OPOL) practice, where each parent has a different native language and speaks that language to the child. One condition for participating in the study was that one of the parents was a native French speaker and the other a native English speaker. All the parents had at least basic knowledge in the other language. We also tried to control variables such as gender and general sociolinguistic environment. An effort was made to obtain a similar number of female and male children as well as children living in France or in the UK. The reason for ensuring an equal number of children from both
countries is that children living in the UK could be said to be exposed to more English and vice versa. As one of our aims was to look at language exposure, it was important to have children living in both countries. Table 1 shows the distribution of the participants.

As is often the case in child language research, all the families were from a middle-class background. The families were asked to supply data about their child’s health history, and their own education, occupation and other biographical questions. The families were studied for, on average, two years from when the child started to talk (approximately four consistent phonological word stage).

The data for the present study consist of tape-recorded conversations between the children and their parents in both monolingual situations interacting in the parent’s native language, and bilingual situations when both parents were present, parental diary of major interactional events and milestones in the children’s development, and parental reports using adaptations of the MacArthur Communicative Developmental Inventory (CDI). For the CDI data, the British English toddler version adapted by Klee and Harrison (2001) was used. It contained 672 words organised into 22 semantic categories. The French toddler version (FCDI) used was called *Inventaire Français du Developpement Communicatif chez le nourrisson: mots et phrases*, adapted by Kern (2001). It contained 689 words organised in the same 22 semantic categories. Each parent in our study was asked to fill in the CDI monthly. The parents were sent (or given) each month (on the day of the child’s birthday) a copy of the FCDI and ECDI. They were asked to fill them in as the CDIs arrived. Further copies of the CDIs were not sent until the previous ones had been returned. Each parent would fill in the CDI that corresponded to the language he/she spoke to the child. The data collection process stopped when the children reached approximately 500 words in at least one of the languages. We fully recognise the shortcomings of parental reports such as the MacArthur CDI (see David, 2004, for a discussion). Parental reports are based on what parents believe their children might be able to produce at the time. Many studies have shown parents reports are a reliable measure of children’s language use, especially vocabulary use (e.g. Thal, 2003). Nevertheless, it is important to bear in mind that what is reported here is the results of parental reports and not necessarily comparable with results based on records of spontaneous production data. One of our key objectives was to address the question whether bilingual children’s lexical development was the same or different as that of their monolingual counterparts. As both the English and French CDIs have already generated a great deal of published data, it seemed a sensible, although by no means ideal, step forward in order to carry out some comparative analysis.

The completed CDIs were analysed initially separately. They were compared at a later stage in order to calculate the number of cross-linguistic synonyms or translation equivalents (TEs) produced by the children. Both forms of the CDI were matched for synonyms. Most of the matches were obvious as the two forms were adapted from a single one and cover the same aspects of a child’s lexicon. But for other concepts, matches were more complex. For example, the word *ball* in English can have two synonyms in
Table 1 Children’s background information

<table>
<thead>
<tr>
<th>Child’s name</th>
<th>Gender</th>
<th>Country of residence</th>
<th>French-speaking parent</th>
<th>English-speaking parent</th>
<th>Family language</th>
<th>Family ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>THO</td>
<td>Male</td>
<td>France</td>
<td>Father</td>
<td>Mother</td>
<td>French</td>
<td>Only child</td>
</tr>
<tr>
<td>FLO</td>
<td>Female</td>
<td>England</td>
<td>Mother</td>
<td>Father</td>
<td>French</td>
<td>Oldest*</td>
</tr>
<tr>
<td>ANT</td>
<td>Female</td>
<td>England</td>
<td>Mother</td>
<td>Father</td>
<td>English</td>
<td>Only child</td>
</tr>
<tr>
<td>EMM</td>
<td>Female</td>
<td>France</td>
<td>Father</td>
<td>Mother</td>
<td>French</td>
<td>Youngest (of 2)</td>
</tr>
<tr>
<td>DUN</td>
<td>Male</td>
<td>France</td>
<td>Mother</td>
<td>Father</td>
<td>English</td>
<td>Youngest (2)</td>
</tr>
<tr>
<td>PEN</td>
<td>Female</td>
<td>France</td>
<td>Mother</td>
<td>Father</td>
<td>English</td>
<td>Youngest (2)</td>
</tr>
<tr>
<td>ELI</td>
<td>Female</td>
<td>England</td>
<td>Mother</td>
<td>Father</td>
<td>English</td>
<td>Only child</td>
</tr>
<tr>
<td>JAC</td>
<td>Male</td>
<td>France</td>
<td>Father</td>
<td>Mother</td>
<td>English/French</td>
<td>Youngest (2)</td>
</tr>
<tr>
<td>ANN</td>
<td>Female</td>
<td>England</td>
<td>Father</td>
<td>Mother</td>
<td>English</td>
<td>Youngest (2)</td>
</tr>
<tr>
<td>OLR</td>
<td>Male</td>
<td>France</td>
<td>Father</td>
<td>Mother</td>
<td>French</td>
<td>Oldest*</td>
</tr>
<tr>
<td>LIA</td>
<td>Male</td>
<td>France</td>
<td>Father</td>
<td>Mother</td>
<td>French</td>
<td>Only child</td>
</tr>
<tr>
<td>REB</td>
<td>Female</td>
<td>England</td>
<td>Mother</td>
<td>Father</td>
<td>English</td>
<td>Youngest (3)</td>
</tr>
<tr>
<td>OLB</td>
<td>Male</td>
<td>England</td>
<td>Mother</td>
<td>Father</td>
<td>English</td>
<td>Youngest (2)</td>
</tr>
</tbody>
</table>

*Younger sibling born during study
French, *balle* and *ballon*, depending on what type of ball one is talking about. *Balle* refers to something smaller (e.g. a tennis ball) than *ballon* (e.g. a football ball). In those cases, one word could have two synonyms. Overall, what these matches aim at demonstrating is ‘that there is a lexical representation from a given “semantic target region”’ (Pearson, 1992). Onomatopoeias were excluded from these matches. The way cross-linguistic synonyms were calculated is discussed at length in David (2004).

In addition, with each CDI checklist, parents were asked to quantify the language input received by the child during the recording month. At the end of the questionnaires parents were asked to evaluate the percentage of French and English to which the child had been exposed to in that month. These estimates provide the basis for the exposure and environment analysis below. Some of the estimates also took into account information gathered from the first author’s interviews with the parents. For example, if a child had gone on a three-week holiday with his French-speaking grandparents in France, the researcher would estimate that 75% of the input for that month was had been French. Table 2 below shows every single month at which CDIs and input questionnaires were collected.

Finally, a number of hours of naturalistic language samples were obtained for each child. The recordings of interactional data were mostly done by the parents but occasionally by the researcher (first author of this paper) on her visits to the families. A total of more than 40 hours of naturalistic recording were obtained. The results for each child are very different depending on the parents’ willingness. Although, ideally, we would have liked the recordings to be regular (once a week), they were quite irregular for most families,

<table>
<thead>
<tr>
<th>Parent</th>
<th>Month 16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>THO</td>
<td>17</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLO</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>ANT</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMM</td>
<td>16</td>
<td>18</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUN</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>PEN</td>
<td>17</td>
<td>18</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELI</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JAC</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>29</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>ANN</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>OLR</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIA</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>REB</td>
<td>16</td>
<td>17</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>23</td>
<td>24</td>
<td>26</td>
<td>29</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLB</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
depending on their own situation. Some were more regular than others. The amount of oral data available for each child varies enormously as families would spend more or less time recording their child. Recordings were orthographically transcribed using the CHAT and LIDES format of CHILDES (see LIDES coding manual from the LIPPS Group, 2000). The utterances were especially coded for switches/mixes. Recordings of a poor sound quality were excluded at this stage of the process. The recordings were transcribed by the first author, a French–English bilingual. Some recordings were checked for reliability by a couple of other bilingual speakers until full agreement was reached. Table 3 highlights the months at which at least 30 minutes of recordings were made.

The following discussion is based on both the CDI data and recorded interactional data, focusing on individual differences.

Data Analysis

As has been reported elsewhere (David, 2004), one of the most important findings of the present study is that bilingual children’s lexical development is very similar to that of monolingual children in terms of distribution of different semantic categories as specified in the CDIs. The profiling of the bilingual lexicon (as reported in David, 2004; David & Li, 2005) showed that bilingual children’s lexical categories in each language developed in a parallel manner (Figure 1 shows the example of nouns).

All of the children acquired social words before nouns before predicates. Even when there is clear imbalance in language dominance, this general pattern is obtained. No child had a lexicon composed of just verbs in one language and nouns in another, for examples. The overall distribution of lexical items, in percentage terms, across the different categories as measured by the CDIs is identical to that of monolingual’s as reported in the literature (for more details, see David, 2004).

Despite the similarities between bilingual and monolingual children in their lexical development, we found considerable variations across individual children. Table 4 gives an overview of the size of the vocabularies for 11 of the children in our sample at 23 and 30 months old, as reported in the CDIs.

Table 3 Oral naturalistic data

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FLO</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>ANT</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMM</td>
<td>23</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUN</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN</td>
<td>19</td>
<td>20</td>
<td>22</td>
<td>23</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLR</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLB</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>27</td>
<td>28</td>
<td>30</td>
</tr>
</tbody>
</table>
It is obvious that great variability exists across the individuals. At 30 months old, for example, FLO produced 993 lexical items, according to her parents, while DUN produced nine times fewer words than FLO (109). The table also shows that although OLB only had 156 words at 23 months, by 30 months he had caught up with ANN despite the fact that he had three times fewer words than she did seven months earlier. An important implication of this finding is that single case studies, as is often the case in bilingual acquisition research, may in fact present a distorted picture of bilingual children’s language development at specific times.

<table>
<thead>
<tr>
<th>Table 4 Lexicons at 23 and 30 months old</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>23 months</strong></td>
</tr>
<tr>
<td>THO</td>
</tr>
<tr>
<td>FLO</td>
</tr>
<tr>
<td>EMM</td>
</tr>
<tr>
<td>DUN</td>
</tr>
<tr>
<td>ELI</td>
</tr>
<tr>
<td>JAC</td>
</tr>
<tr>
<td>ANN</td>
</tr>
<tr>
<td>OLR</td>
</tr>
<tr>
<td>LIA</td>
</tr>
<tr>
<td>REB</td>
</tr>
<tr>
<td>OLB</td>
</tr>
</tbody>
</table>

Figure 1 Development of nouns in French and English

The International Journal of Bilingual Education and Bilingualism
It was not only vocabulary size that differed significantly from one child to another. The figures in Figure 2 represent two of the children’s lexical development in English and French respectively, as measured by the parental reports. Although we argued earlier that lexical development of bilingual children was very similar to that of monolinguals, it does not mean that the two languages in the developing lexicon progressed in the same way and/or at the same speed. The two languages followed a very parallel development for the child represented in the left-hand chart (OLR). In contrast, the child (PEN) represented in the right-hand chart did not develop her English at the same rate as her French at all. Although the two languages did appear to develop at the same rate in the earliest stages (until 18 months old), later, French developed much more rapidly than English after this point.

Moreover, our study observed very significant differences in the number of translation equivalents produced by the children (as illustrated in Figure 3).
Figure 3 presents the minimum, median and maximum percentages of translation equivalents (or cross-linguistic synonyms) within the children’s average lexicon for each month of the study. It is based on data from both CDIs matched for synonyms. One can see that there is a considerable amount of variability across the children at any given age. Forty percent of what some children are said to be producing are cross-linguistic synonyms. But there are also some children who only know one word for each label (see David, 2004, for more details).

Let us now turn to the extralinguistic factors that may help to explain such important individual differences.

Gender

The first factor to be considered here is gender. The overall average mean of words produced over the whole time of the study is higher for the girls than the boys (387 versus 133), as can be seen in Table 5. Figure 4 presents the central tendency (mean), maximum and minimum for each gender each month.

The figure shows that there is a consistent difference between the boys and the girls, and that the girls have a much larger vocabulary than the boys and that the difference widens as they grow older. The difference between boys and girls is statistically significant. A two-sample t-test reveals that $p = 0.009$. The variation within each gender group is also very important. Girls, especially, show a lot more variance than boys. A test for equal variance suggests that boys and girls variances are statistically different ($p = 0.003$) (Figure 5).

These results are consistent with previously reported findings for monolingual acquisition. Kern (2000) reported a significant difference between boys’ and girls’ production data from the FCDI favouring girls. Le Normand et al. (2002) found similar differences for French acquisition. Fenson et al. (1993) also noted the higher score of girls overall in the American CDI. However, Kern (2001) mentioned that at 30 months old, boys had caught up with girls and produced on average slightly more words than girls. The reason behind this is not clear though. Le Normand et al. (2002) interpret their results as a memory issue. They advance the possibility that girls might have a better declarative memory than boys. This part of the memory is responsible for the mental lexicon and the set of grammar rules children need to build utterances. The fact that the bilingual girls and boys in our study appear to be behaving in the same way as monolingual children is not in itself surprising. Bilingual children would then appear to have and use the same type of memory as monolingual children to process their lexicon and utterances regardless of whether they have one or two languages to deal with.

**Table 5** Mean number of words produced for girls and boys

<table>
<thead>
<tr>
<th></th>
<th><strong>Mean</strong></th>
<th><strong>sd</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td>133</td>
<td>132</td>
</tr>
<tr>
<td>Girls</td>
<td>387</td>
<td>312</td>
</tr>
</tbody>
</table>
We now turn to the factor of language exposure. Our main concern here was to find out whether lexical development would be different if a child had more exposure in one language than another and to quantify the difference. Exposure was measured in terms of the amount of time a parent spent with the child and other environmental (linguistic) factors such as attendance at local nurseries or having a nanny or other caregivers. Parents were asked to

**Figure 4** Gender differences over time

**Language exposure**

Exposure was measured in terms of the amount of time a parent spent with the child and other environmental (linguistic) factors such as attendance at local nurseries or having a nanny or other caregivers. Parents were asked to

**Figure 5** Test of equal variances for girls and boys
answer a few questions each time they completed a CDI estimating the amount of time the child might have spent with speakers of each language over the past month. The estimate was converted to a percentage. So, parents would provide two CDI scores (one for English and one for French) as well as a percentage of exposure to French and English for that month.

Statistical analysis shows that there was a significant correlation between language exposure and vocabulary size ($r = 0.650$, $p < 0.01$, $N = 131$). Therefore, the more a child was exposed to French, the more lexical items s/he was likely to have in that language. Children with a more balanced input (between 40% and 60%) appeared to have the most balanced lexicon (Figure 6).

One related question here is whether major changes in the amount of language exposure had any effect on the child’s vocabulary size. What we meant by changing environment is a change in language exposure (as reported by the parental reports and the researcher’s observations) greater than, or equal to, 20% from one month to the next. So, for example, if a child goes away to spend three weeks with her grandparents who only use one language, her language exposure would change dramatically for that month compared to the previous and following months. The results show that there is a small significant correlation between changing environments and lexicon size (i.e. the child’s CDI score in the corresponding language) ($r = 0.246$, $p < 0.01$, $N = 131$). This correlation was calculated for every single month for every child for whom data was available between the percentage of exposure to French and the percentage of the French lexicon in the total lexicon. It is interesting to note that two children with the largest vocabulary were children whose input varied the least. However, the two children with the smallest did not experience major changes in the amount of language exposure over time. So, other factors must play a role, some of which will be discussed further below.

An illustration of the effects of changing language exposure can be seen in Figure 7 that represents one child, REB’s, changing exposure to French and the proportion of French lexical items in her total lexicon. The line represents the

![Figure 6](image_url) Proportion of the French lexicon as a function of the exposure to French
child’s proportion of exposure to French in percentage (as opposed to exposure to English). Once again, that was calculated based on the parents’ report and the researcher’s observations and correspondence with the mother. The bars represent the child’s proportion of French items in her total lexicon.1

We can see that as REB’s exposure to French decreases, the proportion of her French lexicon in the total lexicon decreases too. REB lives in an English-speaking environment (England). At the beginning of the study, REB was cared for at home by a native French speaker (a family relative) and her French-speaking mother and by the end of the study was going to an English-speaking nursery while her mother worked gradually more often. Between 19 and 24 months, the stabilisation and even slight increase in the amount of exposure to French seems to result in the slower decrease (or even slight increase) in the proportion of her French vocabulary at 24 months. The sharp decrease in her exposure to French at 26 months is due to a visit to her English-speaking grandmother for three weeks on her own, without her parents.

It seems reasonable to speculate that differences in language exposure not only affect a bilingual child’s vocabulary size but also her knowledge of the lexical items in the two languages. Specifically, we hypothesise that the number of translation equivalents a bilingual child has may depend on the language dominance of the child and language dominance is at least in part a result of language exposure. No study, has, to our knowledge, looked at the link between translation equivalents and language exposure.

A Pearson correlation was calculated in order to measure and quantify the relationship between the number of translation equivalents in a child’s lexicon in any given month and the same month’s exposure to French as described above (in percentage). The result of the correlation is statistically significant (at the 0.05 level), suggesting that translation equivalents are influenced by language exposure ($r = -0.198$, $p = 0.23$, $N = 131$).

We can conclude that the more balanced the child’s language exposure is, the more likely the child is to have cross-linguistic synonyms in his/her lexicon. These results demonstrate that not only does language exposure

---

**Figure 7** Effect of change in language exposure on vocabulary size
influence the quantity of lexical items acquired in each language but also affects the nature of the bilingual lexicon. The presence or not of translation equivalents depends significantly on language exposure and in turn on language dominance.

**Parental language**

We now consider the relationship between bilingual children’s lexical development and their parents’ input. We do so by looking at both quantitative measures of mean length of utterance (MLU) and lexical diversity (measured here with TTR) in the parental speech and qualitative measures of parental discourse strategies. These measures were obtained by using CLAN. The data of parental speech comes from recordings of parent–child interaction. It is important to note that the use of MLU and Type Token Ratio (TTR) measures is widely used in child language acquisition with relatively reliable results up until a certain stage. However, the use of similar measures for older children and adults poses certain problems that have lead Dewaele (2000), for example, to claim that they are unreliable for adult population. For example, in naturally occurring conversations with adults, very short sentences (e.g. *Oh yes!* ) tend to bring the average of MLU to a very low score, which in reality does not reflect the adults’ linguistic abilities. This claim can be further illustrated by the fact that the standard deviation in the parents’ MLU (*sd* = 0.9) is wider than that of the children’s (*sd* = 0.7). MLU and TTR scores for parents in the present study are not in any sense an exact measurement of their capacities. It is merely a readily available tool that enables a correlation between parents’ language and their children’s lexical development.

Table 6 highlights the different correlations between the child’s and his/her parents’ speech. No distinction has been made between languages in this instance so as to provide an overall picture.

It appears that the child’s lexical development is closely linked to the parents’ MLU and lexical diversity in their speech. Parents who use longer and more complex utterances and a wider lexical range contribute to children’s bigger vocabulary size. And parents tend to increase their utterance length and lexical diversity as the child’s abilities increase. One could hypothesise that parents notice the child’s growing abilities and therefore use more complex and longer sentences. Or it may be a case that the child’s abilities impact on the parents’ speech. It is likely that it is a ‘virtuous circle’. Both influence each other and the child benefits from the whole process. It allows him/her to improve his/her language skills. Further results confirm a positive correlation (*r* = 0.504, *p* < 0.01, *N* = 57) between the children’s amount of

<table>
<thead>
<tr>
<th></th>
<th>Child’s lexicon</th>
<th>Child MLU</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU parent</td>
<td>0.630**</td>
<td>0.621**</td>
</tr>
<tr>
<td>TTR parent</td>
<td>0.485**</td>
<td>0.515**</td>
</tr>
</tbody>
</table>

**Table 6** Correlations between parental speech and children’s lexical development

**Significant at *p* < 0.01**
cross-linguistic equivalents produced and the parents' lexical diversity measure.

With regard to parental discourse strategies, Döpke (1992) found that the way in which parents applied the one person–one language (OPOL) strategy affected the way the children used their languages. If the parents applied a strict OPOL rule, then the child had clearer cues to understand the bilingual situation s/he was in and was therefore helped to develop cross-linguistic synonyms. In order to test that quantitatively, we correlated the number of TEs a child would have in his total vocabulary at the time of a given recording with the amount of parental codeswitching (proportion of utterances containing words belonging to more than one language). However, we did not find that the amount of codeswitching done by the parents influenced the number of TEs produced by the child or that the type of parental strategies, in case of mixing from the child, correlated with the number of TEs. The correlations were not significant at the individual level or at the group level. That is not what might have been expected in light of Döpke’s claims (1992). Nonetheless, our sample might have been too small to show a significant correlation.

d. Birth-order/siblings

The last variable we consider is birth order or the presence of siblings. Many studies of monolingual children (e.g. Kern, 2001; Maital et al., 2000) have shown the effects of birth order on language development. However, no correlation was found between vocabulary size and birth order in the present study. This could be due to the relatively small sample size.

Summary and Conclusion

We have reported in this paper findings from a longitudinal group study of lexical development of French–English bilingual children. We have provided evidence of significant individual differences in the children’s language development, an issue that is impossible to study with single cases. We have examined some of the factors which may have contributed to the differences in vocabulary sizes, number of cross-linguistic synonyms and amount of language mixing. First, we have shown a significant difference between mean vocabulary sizes of boys and girls. Girls appear to have a larger vocabulary than boys and develop quicker. That difference does seem to fade away by 30 months, as it has sometimes been claimed in monolingual studies. Our finding is consistent with some studies which reported a superiority of girls in early lexical development (e.g. Le Normand et al., 2002).

Most importantly perhaps, parental input plays a crucial role in children’s lexical development. The amount of language exposure in each language significantly affects the children’s vocabulary size and language dominance. In addition, the way parents talk to their child influences the size of their lexicon. That is also consistent with studies of monolingual development which claimed that quality and quantity of input do matter in lexical acquisition (e.g. Hart & Risley, 1995).
In terms of cross-linguistic synonyms, we have suggested that the richness of the vocabulary employed by the parents and how balanced the child is in the two languages lead to individual differences. Whilst no direct evidence of parental strategies influencing the acquisition of TEs was observed in quantitative terms, the example of FLO and her parents’ use of the ‘steering strategy’ have shown that translation equivalents and parental strategies are linked.3

One finding of our study that differs from those of other studies of monolingual children is the birth-order effect. We did not find a correlation between birth-order and lexical development. However, this may be due to the relatively small sample we have in the study.

We hope to have brought further evidence to the fact that bilingual children develop in a similar way to monolingual children. Factors influencing lexical developments in monolinguals have similar consequences in bilinguals. Our findings integrate into the current trend of thought that bilingual language development is by and large similar to monolingual language development. As far as the cross-linguistic influences debate is concerned, this study has also shown that developing in the same way as monolinguals does not necessarily mean that a bilingual child’s two languages develop at the same time. Lexical data presented here clearly illustrate the fact that most of the children acquired one language quicker than another one. Most of the children were clearly dominant (in lexical terms) in one language at one given point in time. The literature often uses OPOL children as subjects in their case studies. These children are said to be more balanced than children growing up in one language–one environment settings, for example. Parents and researchers often use or choose that method in the hope that the child will be getting a more balanced input than in other situations. However, this study has shown that this is not always the case. The OPOL strategy does not always mean more balanced input and lexicon. Consequently, it is very important for researchers to quantify and/or qualify language exposure and input whenever working with bilinguals.

The present study highlights the importance of individual differences in bilingual acquisition. To investigate individual differences seriously, group studies are clearly needed. Such studies require more resources than single case studies and can therefore be more demanding. Yet a fuller understanding of bilingual development cannot be achieved without an understanding of individual differences. It is hoped that the present study acts as a catalyst for future comparative research on bilingual children.

Acknowledgements

This paper is based on findings of the first author’s PhD project, supervised by the second author, which was part funded by a grant of the Economic and Social Research Council of the UK. Some of the findings were presented at various conferences and seminars. We are grateful to Margaret Deuchar for her invitation to the seminar and her encouragement to write up the paper, and to all the participants for their feedback. We thank Jean-Marc Dewaele, who
examined the thesis and made detailed comments on the study. We also thank two anonymous reviewers for their most constructive comments on the paper.

**Correspondence**

Any correspondence should be directed to Prof. Li Wei, School of Languages, Linguistics & Culture, Birkbeck, University of London, London, UK (li.wei@sllc.bbk.ac.uk).

**Notes**

1. Total lexicon = sum of French lexicon + English lexicon.
2. As a reviewer pointed out, it is difficult to know who affects who without further analyses.
3. This is an example where a case study of that child in particular helps to show aspects of language development which might not otherwise be visible in larger group studies. Our group study is, nonetheless, small enough to enable case studies within the larger group study.
4. It is important to note here that throughout this paper, children have been matched for age. Although it can be argued (as pointed out by a reviewer) that MLU-match may be more appropriate as it is well documented that children develop at different rates, this was not possible in the present study. Only two children were recorded until their MLU reached 4. Therefore, results would be biased as those two children had a very large lexicon.

**References**


