Parental reports of bilingual toddlers’ first words and grammar of Italian in majority and heritage language contexts

Francesca La Morgia and Jo Billington
Trinity College Dublin and University of Reading
flamorgi@tcd.ie

Abstract
This study presents the results of an investigation into the size and composition of vocabulary and early sentence formation in two groups of young bilingual children, one acquiring Italian as a majority language and the other acquiring it as a heritage language. The results show high variability in the relationship between input and vocabulary size and composition in children between the ages of 24 and 29 months, while this relationship becomes more stable in children between the ages of 30 and 37 months. The results also show that the majority of children have a larger vocabulary and produce more complex sentences in the majority language, input received in each language does not systematically correlate with vocabulary size or with production of complex grammatical structures. The results shed light on some characteristics of early simultaneous bilingual language development. Implications for practice are discussed.

Keywords: language acquisition; bilingualism; heritage language; vocabulary; input

1. Introduction
In heritage language (HL) acquisition, children are exposed to the majority language and to a minority language (the HL), and the two languages are normally used for different purposes and hold a different status in society. The majority language usually becomes the dominant language at some point during childhood, and HL speakers tend to underperform in the HL when their scores in certain tasks are compared to those of monolinguals (Pascual y Cabo & Rothman 2012). As Hoff, Welsh, Place, & Ribot (2014) observe, bilingual children hear each language less than the average monolingual child, and the great variability in language skills
found among bilinguals can be attributed to their language exposure as well as their opportunity to use the language. It is therefore important to consider environmental variables when studying or assessing the performance of HL speakers. An environmental variable that has been widely investigated is the quantity of input, which has been found to correlate with language dominance and attainment in various linguistic domains (Hoff, Core, Place, Rumiche, Señor, & Parra 2012; La Morgia 2015; Treffers Daller & Silva Corvalán 2015; Unsworth 2014), and particularly with lexical development (Dixon 2011; Gathercole, Thomas, & Hughes 2008; Hurtado, Grüter, Marchman, & Fernald 2014; Scheele, Leseman, & Mayo 2010). Other studies suggesting that there is an effect of the linguistic experience on the child’s vocabulary development have shown that children who have two parents who are speakers of the minority language are more likely to develop a richer vocabulary in that language than those who only have one parent who speaks it, and they are also more likely to maintain the language over time (Hoff, Rumiche, Burridge, Ribot, & Welsh 2014).

Another important aspect that needs to be taken into account when studying vocabulary in relation to language exposure is that the process of learning a word requires children to hear and use the word in context (Hart & Risley 1995), and that bilingual children hear each language less than the average monolingual child (Hoff, Welsh, Place, & Ribot 2014); therefore the comparison with monolingual norms will in most cases put the heritage language speaker at a disadvantage. In fact, research on bilingual toddlers has shown that the development of their expressive vocabulary is comparable to that of monolingual children; however, where only one language is accounted for, some studies found bilingual children to have smaller vocabularies compared to monolinguals of the same age (Dixon, Wu, & Daraghmeh 2012; Hoff et al. 2012; Pearson, Fernández, & Oller 1993; Thordadottir, Rothenberg, Rivard, & Naves 2006), while some have found that there are no significant differences (De Houwer 2010; De Houwer, Bornstein, & Putnick 2013; Smithson, Paradis, & Nicoladis 2014). However, when both languages are tested, and the total conceptual vocabulary is measured, bilingual children have consistently been shown not to lag behind monolinguals (Thordadottir, Rothenberg, Rivard & Naves 2006). It is therefore important to take into account the context of acquisition, but also to use measures that allow for comparison with other bilinguals rather than just with monolingual children.

One of the instruments that have been widely used for the measurement of the size and composition of young children’s vocabularies is the MacArthur-Bates Communicative Development Inventory (MB-CDI). The CDI consists of a checklist of children’s lexical and
morpho-syntactic competencies between the age of 8 and 36 months, and has been employed in monolingual, bilingual and cross-linguistic studies, as it is available in several languages (Devescovi, Caselli, Marchione, Pasqualetti, Reilly, & Bates 2005). Its validity has been widely documented and it has been shown that parents’ reports reflect the children’s abilities demonstrated in spontaneous language production (Devescovi & Caselli 2001; Fenson, Marchman, Thal, Dale Reznick, & Bates 2007; Law & Roy 2008; Marchman, Martinez-Sussman, & Dale 2004; Paradis, Emmerzael, Sorenson & Duncan 2010; Pearson & Fernandez 1994).

The Italian CDI, Parole e Frasi nel Primo Vocabolario del Bambino (Caselli & Casadio 1995; Caselli, Pasqualetti, & Stefani 2007) was developed in parallel to the English version (Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly 1993), and it was designed to allow cross-linguistic comparison. The results of the crosslinguistic studies based on the data collected from monolingual Italian and monolingual English speaking children have demonstrated that the two languages are comparable in terms of onset and development of lexical categories, relationship between vocabulary size and grammatical complexity, and order of acquisition of function words. The main difference was found in the larger amount of morphology acquired by Italian children. This was explained by the rich system of morphology that characterizes Italian, in comparison to English, which makes children more sensitive to verb agreement at a very young age (Caselli, Casadio, & Bates 1999).

1.1. This study
The present study examines the size and composition of the vocabulary and the early sentence production in two groups of children, one acquiring Italian as a majority language and one acquiring it as a HL. Both groups only include simultaneous bilingual children, therefore avoiding the comparison with monolinguals. The aim of the study is to determine whether there is a link between exposure to the HL and CDI scores in young children who are at the early stages of language production. It is expected that some patterns in the scores will be common to both languages spoken by the child, by some may only apply to one of the languages. By pointing out some features that are typical of the HL, this study will endeavour to contribute to our understanding of the nature of simultaneous bilingual development.

Previous studies have used parents’ reports to examine the relationship between input and expressive vocabulary in other languages pairs (Pettenati, Vacchini, Stefanini, & Caselli 2011) and cross-linguistic differences between Italian and English monolingual children, but
no study so far has examined the vocabulary and grammar of bilingual children acquiring English and Italian simultaneously in relation to language exposure. The design of this study is based on Pettenati et al.’s (2011) and the original data from Italian-English bilingual children collected for this study will be presented in a way to mirror the findings from Pettenati et al.’s work, in order to present similarities and differences between the two groups of bilingual children.

2. Method

2.1. Participants

This study presents results from an original dataset consisting of scores from the Italian and the English version of the CDI and a questionnaire on language input and language history compiled by 16 Italian-English bilingual families, and compares these to the findings from the study by Pettenati et al. (2011).

The data collection for this study was carried out in England and Ireland, where English is the majority language, and Italian is acquired as a HL. Sixteen bilingual Italian-English families participated in this study. All children were simultaneous bilinguals born and raised in England or Ireland. All children (eight males and eight females) had at least one Italian parent, and they were exposed solely to Italian and English from birth. Six children had two Italian speaking parents, 10 children had one Italian and one English speaking parent. Out of the 16 children, 10 were in English-speaking daycare for an average of 6 hours a day, while the others were cared for at home by family members. The children’s ages ranged from 24 to 37 months. Twenty-four months was chosen as the lowest age, following guidelines from other studies (Pettenati et al. 2011; Gatt, O’Toole & Haman 2015), which are based on the evidence that across different languages children have been found to use over 150 words at 24 months (Bleses et al. 2008). Gatt et al. (2015) also suggest that the upper age range when using the CDI should go beyond the 30-month limit recommended by Fenson et al. (1993). Following Pettenati et al. (2011) and Gatt et al. (2015), the age of 37 months was used in this analysis as the upper age range, in order to observe possible age-related changes happening at a very crucial time for language development.

The study by Pettenati et al. (2011) included a similar sample consisting of 12 children (six males and six females) between 24 and 37 months, who were raised in Italy, and were exposed to Italian and Spanish from birth, and attended Italian-speaking daycare for an average of 8 hours a day.
2.2 Materials and procedures

The two versions of the MacArthur-Bates Communicative Development Inventory (CDI) employed in this study were the Lincoln University Babylab Toddler Communicative Development Inventory for the English language and Il primo vocabolario del bambino, Scheda Parole e Frasi for the Italian language (Caselli & Casadio 1995). Parents were asked to fill the CDI in their native language. The six families in which both parents were Italian were asked to seek help in the completion of the English CDI by involving other English-speaking caregivers in the completion of the form, where possible, following De Houwer et al.’s (2013) recommendation. For the purpose of this study, only two parts of the CDI were selected, following Pettenati et al. (2011). The first was the vocabulary section (Lista di parole, in the Italian version) which consists of words from all categories. Parents tick the word that the child says, or replace the word with the form used by the child, if different from the one stated on the list. The Italian list includes 670 items, while the English one includes 695 items, therefore the scores are shown both in numbers and percentages to account for this difference. The other section analysed was the one labelled complexity. This section consists of 37 pairs of sentences; each pair contains one version that resembles telegraphic speech, or represents a simpler version of the other sentence, which includes elements such as determiners, auxiliary verbs and other function words, as well as inflected verbs and subordinate clauses. Parents were asked to select the sentences that are most similar to those produced by their child. The analysis of the scores from the word list and the complexity section was compared to that of Pettenati et al. (2011), in order to examine the differences between children acquiring Italian as a majority language in Italy to those who acquire Italian as a HL in English speaking contexts.

The other tool used was the Language Exposure Calculator, developed for this study on the basis of other existing tools used to measure language exposure in bilingual contexts (Li, Sepanski, & Zhao 2006). This questionnaire was administered to parents at the same time as the CDI, and required them to provide the parents’ and the child’s biographical information, language history and current use and sources of language exposure outside the home environment. Both parents also completed a detailed weekly schedule of the child, with a breakdown of activities and interlocutors interacting with the child hour by hour. For each of the child’s daily activities, parents were asked to state who is normally with the child, and what language(s) are spoken by the people involved.
Finally, unlike in Pettenati et al.’s study, for the purpose of this analysis, children were divided into two groups, one including children ranging from the age of 24 to 29 months, and the other including children from 30 to 37 months. This grouping was motivated by some interesting differences and similarities found in the levels of language exposure and the language abilities of the children across the two groups.

3. Results

3.1. Vocabulary composition

The vocabulary composition section of the CDI includes the three categories of nouns, verbs and adjectives. The scores presented in these areas show that the Italian lexical repertoire of the Italian-English bilingual children was made of 35% nouns, 15% predicates, 41% social words and 9% function words; their English repertoire included an average of 58% nouns, 18% verbs, 18% social words, and 6% function words. The Italian-Spanish bilingual children were reported to produce an average of 52% nouns, 23% predicates, 9% function words and 14% social words, while their Spanish repertoire included an average of 62% nouns, 9% predicates, 7% function words and 22% social words. This result is quite interesting, as it shows that the largest category is that of nouns, as expected, but the second largest is different across the two groups. In both groups, the second largest category in the majority language is that of predicates, while in the HL it is that of social words.

A similar pattern was found in the analysis of verbs. In the CDI, these are divided into the categories of action words, consisting of lexical verbs and helping verbs, consisting of modal and auxiliary verbs. The Italian-English bilinguals were found to produce on average 62% of lexical verbs, and 21% of modal verbs in English, while they produced 47% of lexical verbs and 7% of modal verbs in Italian. Within the category helping verbs, children produced an average of 5 out of 21 verbs in English, and an average of 1 out of 20 in Italian. It must be noted that 14 children were reported to produce no modal verbs in Italian, and 8 of those children produce no modal verbs in English. The analysis of the category descriptive words (which includes 63 adjectives in Italian and 64 in English) also reveals that children produced on average 55% of English adjectives, and only 27% of Italian adjectives. As these results show, the group is very heterogeneous, but there is a visible trend showing that the English vocabulary is on average larger, and only 4 children were reported to have a larger Italian vocabulary. Table 1 shows the results from the two groups of children, presenting their vocabulary size as well as the exposure to each language.
3.2. Vocabulary size and language exposure

Across all categories, on average, children from both groups were found to produce more words in the majority language: the English-Italian bilinguals produced an average of 217.8 Italian words (SD=135.8; range: 40–467) and 307.9 English words (SD=199.3; range: 27–596), while the Spanish-Italian bilinguals produced an average of 337.2 Italian words (SD=144; range: 42–582) and 100 Spanish words (SD=93.3; range: 9–293).

As Table 1 shows, all bilingual children aged 30 to 37 months were exposed more frequently to the majority language, with the exclusion of one child who has equal exposure to Spanish and Italian. The results from the younger group exhibit greater variation: for example, while the younger Spanish-Italian bilinguals are exposed more frequently to Italian, their English-Italian counterparts display more variability, with some being exposed to Italian 5% of the time, and some 95%. The overall analysis of the vocabulary size shows that on average children have a larger vocabulary in the majority language. When the groups are sub-divided into age groups, as shown in table 1, more marked differences can be observed. The younger Italian-English bilingual children have on average a larger Italian vocabulary (Italian: M=171.5; SD=136.4; English: M=122.5; SD=171.1), while the older ones have on average a larger English vocabulary (Italian: M 319.6, SD 101.3; English M 495.8, SD 107.5). The analysis of the amount of input shows that the younger children are exposed to Italian slightly more frequently than to English (on average 52.7% of the time), while the older children are only exposed to Italian on average 20% of the time.

The Spanish-Italian bilingual children consistently show a larger Italian vocabulary across age groups: the younger group produce an average of 261 Italian words (SD=122.3) and 139 Spanish words (SD=118.3), and the older group produce an average of 444 Italian words (SD 101.3) and 103 Spanish words (SD 112.8). Unlike the English-Italian bilingual children, the Spanish-English bilinguals from both groups are exposed to Italian more than to Spanish (older group 27% and younger group 32% of the time), and this might explain why their Italian vocabulary is consistently larger.
Table 1: Input and vocabulary size across the two groups

<table>
<thead>
<tr>
<th>Age (mont hs)</th>
<th>Percent age of Italian input</th>
<th>Percent age of English input</th>
<th>Italian vocabulary</th>
<th>English vocabulary</th>
<th>Age (mont hs)</th>
<th>Percent age of Italian input</th>
<th>Percent age of Spanish input</th>
<th>Italian vocabulary</th>
<th>Spanish vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>24-29 months</td>
<td>24</td>
<td>95</td>
<td>5</td>
<td>47</td>
<td>24</td>
<td>70</td>
<td>30</td>
<td>245</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>15</td>
<td>75</td>
<td>147</td>
<td>24</td>
<td>75</td>
<td>25</td>
<td>400</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>5</td>
<td>95</td>
<td>40</td>
<td>24</td>
<td>65</td>
<td>35</td>
<td>182</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>90</td>
<td>10</td>
<td>112</td>
<td>25</td>
<td>65</td>
<td>35</td>
<td>42</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>10</td>
<td>90</td>
<td>467</td>
<td>27</td>
<td>55</td>
<td>45</td>
<td>268</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>85</td>
<td>15</td>
<td>28</td>
<td>28</td>
<td>70</td>
<td>30</td>
<td>375</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>40</td>
<td>60</td>
<td>227</td>
<td>29</td>
<td>75</td>
<td>25</td>
<td>316</td>
<td>329</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>60</td>
<td>40</td>
<td>158</td>
<td>29</td>
<td>75</td>
<td>25</td>
<td>316</td>
<td>329</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>60</td>
<td>40</td>
<td>360</td>
<td>29</td>
<td>75</td>
<td>25</td>
<td>316</td>
<td>329</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>95</td>
<td>5</td>
<td>122</td>
<td>29</td>
<td>75</td>
<td>25</td>
<td>316</td>
<td>329</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>25</td>
<td>75</td>
<td>179</td>
<td>29</td>
<td>75</td>
<td>25</td>
<td>316</td>
<td>329</td>
</tr>
<tr>
<td>30-37 months</td>
<td>30</td>
<td>10</td>
<td>90</td>
<td>369</td>
<td>33</td>
<td>75</td>
<td>25</td>
<td>399</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>20</td>
<td>80</td>
<td>386</td>
<td>34</td>
<td>70</td>
<td>30</td>
<td>306</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>30</td>
<td>70</td>
<td>282</td>
<td>36</td>
<td>50</td>
<td>50</td>
<td>468</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>20</td>
<td>80</td>
<td>326</td>
<td>37</td>
<td>80</td>
<td>20</td>
<td>582</td>
<td>293</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>20</td>
<td>80</td>
<td>235</td>
<td>37</td>
<td>90</td>
<td>10</td>
<td>464</td>
<td>9</td>
</tr>
</tbody>
</table>

3.3 Sentence complexity

The results from the analysis of the section *Complexity* mirror that of vocabulary. The Italian-English bilingual children produced on average 7.1 complex sentences in Italian, and 11.3 in English (out of a total of 37 in each language). Of the 10 children in the younger group, 7 scored 0 in the complexity section of the Italian checklist, and 3 of these children were reported not to produce complex sentences in either language.

All the older children were reported to produce more English than Italian complex sentences (on average, 17.4 sentences in Italian and 24.8 sentences in English). A similar picture emerges from the analysis of the results from the Italian-Spanish bilingual children, all of which were reported to produce more complex sentences in Italian. The difference is more marked than that found for the Italian-English bilingual children, as this group produced on average 29.4 Italian and 7 Spanish complex sentences. Similarly to the findings from the analysis of the vocabulary size, the Spanish-Italian children performed better in Italian than in
Parental reports of bilingual toddlers’ first words and grammar of Italian in majority and heritage language contexts

Spanish, and the gap between the two languages is wider on average than that of the Italian-English children.

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Complexity, Italian</th>
<th>Complexity, English</th>
<th>Complexity, Spanish</th>
<th>Complexity, Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>24-29 months</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>0</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>35</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>12</td>
<td>26</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>5</td>
<td>8</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>13</td>
<td>22</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>22</td>
<td>31</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 2: Number of complex sentences

4. Discussion
The primary purpose of this study was to investigate the size and composition of vocabularies and early sentence formation of bilingual children acquiring Italian alongside another language in two different contexts, where Italian is either a majority language (in Italy) or a heritage language (in Ireland and in the UK). Mirroring the analysis presented in a previous study on Spanish-Italian bilingual children (Pettenati et al. 2011), this study addressed questions related to input and early linguistic development of bilingual children between 24 and 37 months.

The two groups of bilingual children differed in the quantity of exposure to the two languages: while the Italian-Spanish bilingual children living in Italy were exposed to either an equal amount of input or a larger amount of Italian input, 6 of the 16 Italian-English bilingual children were exposed to Italian more often than to English. All but one of the
bilingual children living in Italy had larger Italian vocabularies. However, the bilingual children living in Ireland and the UK showed more mixed results, whereby four of the children who were exposed to Italian more had smaller Italian than English vocabularies. Of these children, three were exposed to English between 5 and 10% of the time, and yet their English vocabulary was larger. While it may be hard to provide an exhaustive explanation for this result given the small amount of data, the data from the 30-37 month group shows more consistency in the relationship between input and vocabulary size across the two groups. Both groups of 30 to 37 month old children are exposed to the majority language more frequently, and all of the majority language vocabulary scores for those children are higher (on average by 258 points, ranging from 40 to 455). These results indicate that the variation found in the vocabulary sizes of 24 to 29 month old children may still be attributed to developmental factors, and that the effects of language exposure may start to emerge after 30 months.

In terms of vocabulary composition, an interesting pattern was found: as expected, most of the words children produced fall into the category of nouns, however, the second largest category is that of predicates only in the majority language, across both groups of children. The second largest category in the HL is that of social words. The explanation for this phenomenon may be the fact that verbs may be acquired more slowly in the HL than in the majority language. This hypothesis is corroborated by the other analysis carried out in this study, namely that of sentence complexity. Both groups display a similar pattern in the use of complex sentences in the HL, with only one child producing more complex sentences in the HL than in the majority language. Children in the older group were found to produce more complex sentences than those in the younger group, but again across the two groups the average of complex sentences in the HL was 12.4, which is a third of the total number of sentences.

These preliminary results gathered using a small data-set constitute a starting point in the investigation of the performance of bilingual children acquiring Italian in different linguistic contexts and they offer results based on data from bilingual children rather than comparisons with monolingual ones. The results have implications for the assessment of bilinguals in clinical settings, as they show that examining only the HL may not be a reliable indicator of the child’s linguistic ability. In addition, the results from this study, which mirror those of Pettenati’s and other studies carried out with simultaneous bilingual children have important practical implications for parents and early childhood educators. As this study shows, the majority language tends to be dominant, even in children who have two parents who speak
the HL. As simultaneous bilingualism does not mean equal knowledge of the two languages, parents and educators should consider the imbalance between the two languages as a normal phenomenon in early bilingual acquisition. Early childhood educators working with simultaneous bilingual children whose home language is different from the country’s majority language should advice parents to provide as much input in the HL as possible, to support the child’s harmonious bilingual development.

References


