ABSTRACT

We present a study examining cognitive functions in late non-balanced bilinguals with different levels of second language proficiency. We examined in two experiments a total of 193 mono- and bilingual university students. We assessed different aspects of attention (sustained, selective and attentional switching), verbal fluency (letter and category) as well as picture-word association as a measure of language proficiency. In Experiment 2 we also compared students in their first/initial (Y1) and fourth/final (Y4) year of either language or literature studies. There were no differences between both groups in category fluency. In selective attention, bilinguals outperformed monolinguals in Y1 and this difference remained significant in Y4 despite overall improvement in both groups. Contrasting results were found in attentional switching and letter fluency: while no differences were found in Y1 in both tasks, in Y4 there was an advantage for bilinguals in attentional switching and for monolinguals in letter fluency. We conclude that overall late-acquisition non-balanced bilinguals experience similar cognitive effects as their early-acquisition balanced counterparts. However, different cognitive effects may appear at different stages of adult second language acquisition.
1. Introduction

1.1. The Cognitive Effects of Bilingualism

Substantial evidence suggests that bilingualism can influence cognitive functions\(^1\). In the linguistic domain, bilinguals show a disadvantage compared to monolinguals in reaction time and accuracy in lexical access tasks such as picture naming\(^2\)-\(^4\), attributed to either parallel activation of words from different languages and the necessity to inhibit competing non-target items\(^5\) or to a reduced-frequency of use of each of the bilingual’s language\(^6, 7\). In contrast, a bilingual advantage has been reported for tests of executive functions, such as attentional control\(^8-12\), inhibition\(^10\) and switching\(^13, 14\). These differences continue across the lifespan\(^12, 15-17\) and might contribute to a later onset of dementia in bilinguals\(^16, 18, 19\). It has been hypothesised that these effects come from higher demands posed on executive control through inhibition and switching between languages associated with bilingualism\(^5\). In some tasks, such as verbal fluency (VF), bilingual performance has shown both advantages and costs. In some category fluency studies, bilinguals have been reported to underperform\(^20-22\), while in others to outperform monolinguals\(^23\). Other authors have reported no influence of bilingualism on category fluency\(^24\). A similar pattern of conflicting results exists in letter fluency\(^20, 24\).

While current debates often focus on the specific nature of the tasks employed\(^13, 14, 25-27\), less attention has been paid to the characteristics of the bilingual speakers and their bilingualism. Most research has been devoted to “classical” bilingualism: a simultaneous or early consecutive childhood acquisition and balanced command of two or more languages. It remains unclear to what extent bilingualism effects can also be detected in individuals who acquire their second language in late childhood or adulthood without reaching native-like proficiency. Studies of late-acquisition bilingualism produced so far conflicting results. Luk et al. (2011) found a bilingual advantage only in early-acquisition bilinguals\(^28\), while other
studies found it in early as well as late-acquisition bilinguals\textsuperscript{17,27,29,30}. Also regarding the importance of the number of languages involved, previous studies came to conflicting results\textsuperscript{31}. Some found a beneficial effect only in multi- but not in bilinguals\textsuperscript{32} or reported a correlation between the number of languages and cognitive performance\textsuperscript{15}. Others found only a weak effect of multilingualism\textsuperscript{17} or no effect at all\textsuperscript{16}.

Against this background, our study set out to examine non-balanced bilinguals who acquired their second language in late childhood/early adulthood. We employed non-verbal auditory tests assessing different aspects of attention\textsuperscript{27} and examined the difference in performance in students in their first/initial and fourth/final year, relating cognitive changes to the increase in L2 proficiency.

\section*{2. Experiment 1}

\subsection*{2.1 Methods}

\subsubsection*{2.1.1. Participants}

Sixty-six University of Edinburgh students (mostly in their 4\textsuperscript{th} year) took part in this experiment. All were native English speakers.

\begin{table}[h]
\centering
\begin{tabular}{lcccccc}
\hline
 & \textbf{Experiment 1} & & & \textbf{Experiment 2} & & \\
 & Monolinguals & Bilinguals & Multilinguals & Monolinguals & Bilinguals & Monolinguals & Bilinguals \\
\hline
\textbf{Total (N)} & 18 & 16 & 17 & 24 & 32 & 22 & 37 \\
\textbf{Age Mean} & 21.78 & 22.44 & 20.82 & 19.67 & 18.75 & 22.09 & 21.70 \\
\textbf{(SD)} & (2.18) & (1.97) & (1.70) & (1.76) & (.67) & (1.11) & (1.37) \\
\textbf{Gender ratio} & 12/6 & 13/3 & 14/3 & 15/9 & 23/9 & 15/7 & 25/12 \\
\hline
\end{tabular}
\caption{Demographic data of the participants.}
\end{table}

The Monolingual participants (N=18) did not speak any language other than English beyond basic level. The Bilingual participants (N=16) had Spanish as their second language (L2) and no knowledge of other languages. The Multilingual participants (N=17) knew at
least one more language in addition to English and Spanish, but their knowledge of Spanish, as indicated in the language questionnaire (Appendix), was better/comparable to that of other foreign language(s). Fourteen participants were excluded because Spanish was not their main L2, one because of incomplete data. Age and gender differences were not significant (chi-square and t-tests all $p > .05$) (Table 1).

2.1.2 Tasks

2.1.2.1 Picture Name Verification Task (PNVT)

The PNVT measures accuracy and speed with which a picture-name combination is judged to be correct or not and provides, therefore, an objective measure of L2 proficiency. The stimuli were 42 pictures depicting clothing, furniture and body parts with corresponding written names in English and Spanish respectively. None of the words were cognates. There was no difference in the number of graphemes between English ($M=5.36$) and Spanish ($M=5.57$) words ($t(41) = -1.013, p > .05$). Colour pictures of the objects were displayed on a white background for 350 ms. before the word appeared next to the image. Both picture and word remained on the screen until the participant responded. The presentation order was randomised. The task was produced and administered using E-prime 2.

2.1.2.2 Test of Everyday Attention (TEA)

The TEA\textsuperscript{33} is a well-established clinical assessment tool, recently applied to measure executive functions in bilinguals\textsuperscript{27}. We selected three subtests, examining different aspects of attention: Elevator Task (ET), Elevator Task with Distraction (ETD) and Elevator Task with Switching (ETS). ET assesses sustained attention: prompted by recording, participants count seven strings of tones, presented at irregular intervals. ETD measures selective attention asking participants to count low tones while ignoring high-pitch ones over ten trials. ETS
requires switching: participants have to use high and low pitch tones as cues for the direction (upwards and downwards, respectively) in which to count ten strings of tones. All tasks were presented through loudspeakers.

2.1.2.3 Verbal Fluency (VF)

The VF tasks consisted of letter and category fluency. Participants were asked to produce as many words as possible within 60 seconds, beginning with the letter F, M and P (letter fluency) or belonging to the category of animals, foods and degree courses (category fluency)\textsuperscript{20, 21, 34, 35}.

2.1.2.4 Language Questionnaire

Participants completed a language questionnaire (Appendix), rating their command of each language in expression, comprehension, reading and writing on a 5-point scale (basic/weak/moderate/advanced/fluent). Total proficiency score was calculated by adding proficiency levels in all domains. The questionnaire was completed after all other tasks.

2.1.3 Statistical Analysis

Analyses of Variance (ANOVA) and independent and related t-tests (as appropriate) were performed to compare mean differences between and within groups. Correlational analyses were conducted using Pearson’s correlation coefficients. Analyses of variables not meeting the assumption of normality were conducted using non-parametric tests. All analyses were performed using SPPS for Windows v.19.

2.2. Results

2.2.1 PNVT
There were no significant differences in *accuracy* to English words between the three groups ($H(2) = .82$, $p = .664$). The bilingual and multilingual groups were significantly less accurate for Spanish than for English words (bilinguals: $z = -2.067$, $p = .039$; multilinguals: $z = -2.217$, $p = .027$), with no difference between bilinguals and multilinguals ($p = .380$) (Table 2).

Table 2. Summary of mean group performance on Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Monolinguals</th>
<th>Bilinguals</th>
<th>Multilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>97.84</td>
<td>98.21</td>
<td>98.32</td>
</tr>
<tr>
<td></td>
<td>(2.97)</td>
<td>(2.95)</td>
<td>(2.35)</td>
</tr>
<tr>
<td>L2</td>
<td>n/a</td>
<td>90.77</td>
<td>94.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(12.83)</td>
<td>(4.90)</td>
</tr>
<tr>
<td><strong>ET</strong></td>
<td>97.62</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>(5.48)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td><strong>ETD</strong></td>
<td>80.00</td>
<td>94.38</td>
<td>94.71</td>
</tr>
<tr>
<td></td>
<td>(22.23)</td>
<td>(11.53)</td>
<td>(8.74)</td>
</tr>
<tr>
<td><strong>ETS</strong></td>
<td>77.22</td>
<td>93.13</td>
<td>82.35</td>
</tr>
<tr>
<td></td>
<td>(22.44)</td>
<td>(10.78)</td>
<td>(21.95)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Monolinguals</th>
<th>Bilinguals</th>
<th>Multilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verbal Fluency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>17.78</td>
<td>17.50</td>
<td>15.47</td>
</tr>
<tr>
<td></td>
<td>(5.47)</td>
<td>(4.55)</td>
<td>(4.46)</td>
</tr>
<tr>
<td>P</td>
<td>16.39</td>
<td>17.44</td>
<td>15.29</td>
</tr>
<tr>
<td></td>
<td>(3.90)</td>
<td>(4.86)</td>
<td>(3.06)</td>
</tr>
<tr>
<td>M</td>
<td>15.50</td>
<td>17.31</td>
<td>15.59</td>
</tr>
<tr>
<td></td>
<td>(4.20)</td>
<td>(4.30)</td>
<td>(3.64)</td>
</tr>
<tr>
<td><strong>Letter Total</strong></td>
<td>49.67</td>
<td>52.25</td>
<td>46.35</td>
</tr>
<tr>
<td></td>
<td>(11.09)</td>
<td>(11.93)</td>
<td>(8.83)</td>
</tr>
<tr>
<td><strong>Animals</strong></td>
<td>25.72</td>
<td>23.94</td>
<td>25.18</td>
</tr>
<tr>
<td></td>
<td>(5.22)</td>
<td>(6.70)</td>
<td>(5.86)</td>
</tr>
<tr>
<td><strong>Food</strong></td>
<td>25.56</td>
<td>25.69</td>
<td>23.82</td>
</tr>
<tr>
<td></td>
<td>(5.61)</td>
<td>(6.36)</td>
<td>(4.31)</td>
</tr>
<tr>
<td><strong>Degrees</strong></td>
<td>21.44</td>
<td>19.44</td>
<td>20.29</td>
</tr>
<tr>
<td></td>
<td>(3.70)</td>
<td>(4.52)</td>
<td>(3.64)</td>
</tr>
<tr>
<td><strong>Category Total</strong></td>
<td>72.72</td>
<td>69.06</td>
<td>69.29</td>
</tr>
<tr>
<td></td>
<td>(12.20)</td>
<td>(15.63)</td>
<td>(11.97)</td>
</tr>
</tbody>
</table>

*Notes:* Accuracy and performance in ET, ETD and ETS are expressed in percentages. For each verbal fluency task, the number of correct words per minute is reported. SD given in parentheses. Significant differences ($p < .05$) are reported on this table as follows: a: ≠ monolinguals, b: ≠ bilinguals, c: ≠ multilinguals.
2.2.2 PNVT in relation to L2 Proficiency

There was a significant positive correlation between self-rated proficiency in Spanish and accuracy to Spanish words in bilingual and multilingual groups, \( r_s = .722, p \ (2\text{-tailed}) < .001 \).

2.2.3 TEA

Prior to analysis, raw scores of the TEA tasks were transformed into percentages. Ninety-four percent of participants performed at ceiling on ET. The few who made an error were monolinguals, but due to the small number of errors the difference failed to reach significance \( (H(2) = 5.73, p = .057) \). A significant group effect was found on ETD \( (H(2) = 9.13, p = .010) \). Pairwise adjusted \( p \)-values comparisons showed that both bilinguals and multilinguals scored higher than monolinguals \( (p = .020 \text{ and } p = .041, \text{ respectively}) \), with no difference between them \( (p > .05) \). On ETS, there was a trend towards a better performance in bi- and multilinguals, but it did not reach significance \( (H(2) = 5.51, p = .064) \).

2.2.5 Verbal Fluency (VF)

No significant differences were found between the three letters or the three categories across groups \( (\text{all } p_s > .05) \) \( (\text{Table 2}) \). More words were produced in category than in letter fluency: monolinguals: \( t (17) = 7.343, p < .001 \); bilinguals: \( t (15) = 5.486, p < .001 \), and multilinguals: \( t (16) = 9.037, p < .001 \), with no differences between the groups in overall score of category or letter fluency \( (p_s > .05) \).

3. Experiment 2

Results from Experiment 1 suggest that late, unbalanced bi/multilinguals performed better than monolinguals on one of the attentional tasks (ETD), showed a trend towards a
better performance on another (ETS) and no differences on VF. Experiment 2 set out to explore these findings in more detail, examining the influence of increased exposure to and proficiency in L2 taking place during language studies. To this end, we compared the performance of first (Y1) and fourth (Y4) year students of Spanish/Italian and of literature/humanities. As we found no significant differences in performance between the Spanish and Italian language groups (all ps > .05), both groups were analysed together. Also, since the bi- and multilingual groups in Experiment 1 did not show major differences, we merged the two groups into one bilingual group. Thus, the focus of Experiment 2 is on the differences in performance between Y1 and Y4 in language and literature students.

3.1 Methods

3.1.1 Participants

A total of 127 first and fourth year students at the University of Edinburgh took part in the experiment. Twelve participants were excluded following the same criteria as in Experiment 1. Age and gender differences between groups were not significant (Table 1).

3.1.2 Tasks

The tasks and procedures were the same as in Experiment 1. A parallel version of PNVT was developed for Italian, containing the same items as the English-Spanish version, but paired with Italian words. Given that no differences were found between the letters and categories in Experiment 1, we reduced the length of our test by restricting it to the letter P and category animals.

3.1.3 Language Questionnaire
Participants completed the same language questionnaire as in Experiment 1, but in addition we also enquired about musical experience (Appendix). No significant differences were found between the groups.

3.1.4. Statistical Analysis

Parametric and non-parametric tests as well as post-hoc pairwise comparisons and correlational analyses were carried out when appropriate. Because of the larger number of participants in this study, between subjects 2x2 ANOVAs with factors group (mono- and bilinguals) and year of study (first and fourth) were carried out to explore possible interactions.

3.2. Results

3.2.1 PNVT

No differences were found between the groups ($F(1, 111) = .010, p = .922, \eta^2 = .000$) or years of study ($F(1, 111) = 3.797, p = .054, \eta^2 = .033$) in the accuracy for English words (a non-significant trend towards improvement occurred in both groups, see Table 3). The bilingual group was more accurate to respond to English (L1) than to L2 words in both Y1 and Y4 (all $p$s $< .002$).

Table 3. Summary of mean group performance on Experiment 2.

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th></th>
<th>Year 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monolinguals</td>
<td>Bilinguals</td>
<td>Monolinguals</td>
<td>Bilinguals</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>97.42</td>
<td>97.55</td>
<td>98.67</td>
<td>98.43</td>
</tr>
<tr>
<td></td>
<td>(3.51)</td>
<td>(3.28)</td>
<td>(1.65)</td>
<td>(2.56)</td>
</tr>
<tr>
<td>L2</td>
<td>n/a</td>
<td>89.86&lt;sup&gt;b&lt;/sup&gt;</td>
<td>n/a</td>
<td>96.24&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(7.12)</td>
<td></td>
<td>(3.81)</td>
</tr>
<tr>
<td><strong>ET</strong></td>
<td>99.40</td>
<td>98.66</td>
<td>98.70</td>
<td>99.23</td>
</tr>
<tr>
<td></td>
<td>(2.92)</td>
<td>(4.23)</td>
<td>(6.09)</td>
<td>(3.27)</td>
</tr>
<tr>
<td><strong>ETD</strong></td>
<td>68.75&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>81.25&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>83.18&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>93.78&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(16.24)</td>
<td>(15.19)</td>
<td>(19.85)</td>
<td>(15.52)</td>
</tr>
</tbody>
</table>
With regards to words in L2, Y4 bilinguals were significantly more accurate ($U = 245.50$, $z = -4.23$, $p < .001$) than Y1 bilinguals (Table 3, Fig. 1).

**Figure 1.** Experiment 2 - Changes in performance between Year 1 and Year 4 on: (a) TEA ETD, (b) TEA ETS, (c) Category Fluency, and (d) Letter Fluency (For the TEA tasks we report the percentage of correct trials, for the verbal fluency tasks, the number of correct words per minute).
3.2.2 PNVT in relation to L2 Proficiency

A significant positive correlation between self-rated L2 proficiency and accuracy to L2 words was found for the bilingual group, \( r_s = .433, p \text{ (2-tailed)} < .001 \).

3.2.3. TEA

No effects or interactions were found on ET (all \( p > .05 \)). On ETD, both groups improved significantly from Y1 to Y4 (\( F(1,111) = 18.406, p < .001, \eta_p^2 = .142 \)), but bilinguals performed better than monolinguals in both years (\( F(1,111) = 13.509, p < .001, \eta_p^2 = .108 \)), with no significant interaction (\( F(1) = .091, p = .763, \eta_p^2 = .001 \)). On ETS, there were main effects of group (\( F(1,111) = 7.797, p = .006, \eta_p^2 = .066 \)) and year of study (\( F(1,111) = 25.491, p < .001, \eta_p^2 = .187 \)), and a significant interaction (\( F(1) = 3.915, p = .050, \eta_p^2 = .034 \)): both groups performed equally in Y1, but by Y4 a significant bilingual advantage was noted (Fig. 1).

3.2.4 Verbal Fluency (VF)

More words were produced in category than letter fluency in all groups (all \( p < .01 \)). With regards to letter fluency, monolinguals produced more words than bilinguals overall (\( F(1,111) = 4.600, p = .034, \eta_p^2 = .040 \)), with a tendency towards significance for the interaction between language group and year (\( F(1,111) = 3.638, p = .059, \eta_p^2 = .032 \)): both groups performed equally in Y1, but a monolingual advantage was observed in Y4 (Fig. 1). In category fluency Y4 students produced more words than Y1 students (\( F(1,111) = 6.528, p < .012, \eta_p^2 = .056 \)), with no differences between the language groups, and no interaction (\( ps > .05 \)).
4. Discussion

Our results suggest that late non-balanced bilinguals experience similar cognitive costs and benefits as their early-acquisition balanced counterparts. A consistent effect across both experiments was a bilingual advantage on ETD, measuring selective attention and, therefore, inhibition of irrelevant stimuli: a task previously reported to be particularly sensitive to late-acquisition bilingualism\(^{27}\). In Experiment 1, there was no additional benefit of multilingualism over bilingualism. If the reason for a bilingual advantage on this task lies in the constant necessity of suppressing the irrelevant language\(^5\), knowing two languages is likely to lead to a ceiling effect, with no further benefit of additional languages. In Experiment 2, the bilingual effect on ETD was already present in Y1 students, in whom the levels of L2 proficiency were relatively modest, and persisted, despite an overall improvement in performance in both groups, into Y4. It is possible that this effect in Y1 can be explained by the fact that some students had previous knowledge of L2 and by the time of testing had completed one term of intensive language study. However, we cannot exclude that superiority on the abilities underlying this test could be a pre-existing cognitive feature predisposing to language studies.

The results on ETS showed a different pattern: all groups performed equally in Y1 but a bilingual advantage appeared in Y4, by which time the bilingual group reached a considerable level of proficiency, as witnessed by significant improvement in accuracy of their L2 responses on PNVT. ETS is a complex task requiring two different processes: inhibition and switching. The latter involves release of inhibition and a potential negative priming effect\(^{36}\), which may be more marked for adult L2 learners, especially in the initial...
stages. The improvement on ETS in Y4 could be linked, therefore, to the higher proficiency in L2 and the increased opportunities for switching between languages.

In VF, an interesting difference was observed between category and letter fluency. In category fluency, no significant differences were found between the mono- and bilingual groups. In contrast, the letter fluency showed a change in performance between Y1 and Y4, not dissimilar to ETS but in the opposite direction. While there was no difference between mono- and bilinguals in Y1, in Y4 the monolinguals outperformed the bilinguals. Since the monolingual group consisted mainly of literature students, this reverse pattern might well reflect four years of intensive engagement with English language in reading, writing and speaking. This finding also suggests that the monolingual participants in our study were comparable in their general cognitive capacity as well as in their academic activities to the bilingual ones. Both language and literature studies showed an improvement in test performance from Y1 to Y4, but it affected different cognitive domains.

Our study has limitations: some students had previous L2 knowledge, so we could not measure their performance at “point zero” of L2 acquisition. We were also not able to compare the same students across their 4-years courses and thus cannot exclude selection biases. However, when designing our study we made a particular effort to minimise potential confounding variables by keeping the sample as homogenous as possible. All participants were students with the same native tongue (English); the L2 was either Spanish or Italian, languages closely related in grammar and vocabulary. In Experiment 2 we were particularly cautious to select the closest possible monolingual control group: students of English literature and humanities from the same university. Both language and literature students had to fulfil the same strict academic criteria in order obtain admission\(^{37}\) and later to progress from the pre-honours (Y1-2) to the honours (Y3-4) stage (interestingly, the percentage of
students who progressed into the honours programme in the three subject areas was practically identical: 92.4% for Spanish, 94.3% for Italian and 92.6% for English). The type of academic activities they engaged in was also broadly comparable, with the main difference being that language students had to read, write, listen and speak in different languages, the literature students mainly in one, English. Accordingly, the greatest improvement for literature students was in letter fluency (specific to English), and for language students in the more general task of attentional switching.

While in some current debates attempts have been made to reduce the effects of bilingualism to a simple difference on a single task, our study emphasises the complex and multidimensional nature of this phenomenon. We suggest that the potential effects of bilingualism on cognition can be positive (e.g. selective attention) as well as negative (e.g. increased speed of lexical access). Some may occur early in the acquisition of L2 or even predate it as a cognitive marker (e.g. ETD), others seem to appear only when reaching considerable levels of L2 proficiency (ETS). More research is needed to explore these differences in more detail. So far, it seems that the cognitive effects of learning L2 in adulthood are not radically different from those of learning one in childhood: a result of considerable interest and relevance to millions of adult L2 learners worldwide.

**Acknowledgements**

We would like to thank Sascha Vriend and Elisa Alongi for the data collection for Experiment 2 and Prof. John MacInnes for his advice regarding the comparability of the different student groups. The first author would also like to thank CONACYT, Mexico and
The University of Edinburgh for supporting this research. A special thank you goes to the editor and three anonymous reviewers for their constructive feedback.

References
