# RaY

Research at the University of York St John

For more information please contact RaY at
ray@yorksj.ac.uk

# The Linguistic and Cognitive Effects of a One-Week Online Italian Language Course: A Pilot Study

**MATTIA ZINGARETTI** ⓘ

**ROBERTA SPELORZI** ⓘ

**TARANJIT K. SOHANPAUL** ⓘ

**ANTONELLA SORACE** ⓘ

*Author affiliations can be found in the back matter of this article

## ABSTRACT

This pilot study examined whether one week of intensive online Italian instruction could yield significant changes in bilinguals' verbal fluency skills in their second language (L2) (Italian) and first language (L1) (English), as well as in attentional skills (selective attention/inhibition and switching). Eighteen undergraduate Italian language students completed verbal fluency tasks in both English and Italian, and subtests of the Test of Everyday Attention, across three timepoints (pre-course, post-course and follow-up). Participants showed a significant increase in L2 verbal fluency following the course, particularly in semantic (category) fluency, with higher L2 proficiency predicting better performance. In contrast, L1 verbal fluency remained stable, and L2 and L1 fluency were positively correlated at post-course and follow-up. Selective attention improved significantly and was maintained at follow-up, while no changes were observed in sustained attention or switching, although the latter was predicted by L2 proficiency and age. Altogether, this pilot study provides preliminary evidence to suggest that even short-term, online L2 learning can yield lasting benefits to both L2 verbal fluency and attentional inhibition.

**CORRESPONDING AUTHOR:**

**Mattia Zingaretti**

Department of Linguistics and English Language, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, Scotland; York St John University, York, United Kingdom

m.zingaretti@yorksj.ac.uk

# 1. INTRODUCTION

## 1.1. THE LINGUISTIC AND COGNITIVE EFFECTS OF BILINGUALISM: MIXED FINDINGS

The linguistic and cognitive consequences of bilingualism – broadly defined as the use of two (or more) languages (or dialects) in everyday life (cf. Grosjean, 2013) – have been the subject of extensive and contentious debate, much of which has stemmed from comparisons between bilinguals and monolinguals. For instance, some studies report that bilinguals exhibit reduced performance in lexical access tasks, including slower reaction times and lower accuracy during word retrieval (e.g., Gollan et al., 2002; Gollan et al., 2005; Ivanova & Costa, 2008). These effects are possibly due to increased competition between languages (Green, 1998) or reduced frequency of use of each language relative to monolinguals, which gives rise to weaker links between lexical forms and their conceptual representations (Michael & Gollan, 2005). At the same time, bilinguals have been found to outperform monolinguals in domains of executive function, such as attention, inhibition and switching (e.g., Bialystok, 1999; Costa et al., 2008; Hernández et al., 2013). While some studies have found cognitive advantages across the lifespan, from childhood to old age, and even a link between bilingualism and the delayed onset of dementia in later life (e.g., Bialystok et al., 2007; Kavé et al., 2008; Alladi et al., 2013), other studies have reported no clear cognitive benefits for bilinguals (Paap & Greenberg, 2013), and concerns have been raised regarding a publication bias favouring studies that report bilingual advantages (de Bruin et al., 2015).

### 1.1.1. Enhancements in executive functions extend to late bilingualism

Despite ongoing debate around the reliability and generalisability of a bilingual advantage, a growing body of research highlights the fact that, where positive cognitive effects of bilingualism are observed, they are not limited to early bilinguals, but can extend to bilinguals who have acquired a second language (L2) later in life (Tao et al., 2011; Vega-Mendoza et al., 2015; Bak et al., 2016). For instance, with regards to improvements in executive functions, Vega-Mendoza et al. (2015) found that fourth-year university language students outperformed their first-year counterparts in attentional switching, suggesting that even classroom-based language learning can yield cognitive benefits comparable to those found in adult bilingualism, when considerable levels of L2 proficiency are reached. Complementing this, Bak et al. (2016) found that just one week of intensive Gaelic instruction led to significant improvements in attentional switching in adults of different age groups (18–78 years), including older adults. These effects remained stable nine months later, but only in those participants who continued to practice Gaelic for at least five hours per week. In particular, improvements were significantly greater than in both passive and active control groups, indicating that the linguistic component of the learning experience, rather than general cognitive stimulation, was key.

### 1.1.2. The other side of the coin: First language attrition in late bilingualism

Alongside the above-mentioned changes in executive functions, bilingualism research has highlighted that acquiring an L2 can lead to a range of changes in the first language (L1) collectively referred to as 'attrition', encompassing temporary changes in online processing and more permanent changes in underlying linguistic knowledge (Schmid & Köpke 2017a, 2017b; Schmid & Köpke, 2019). Importantly, attrition has been observed not only in immersed bilinguals but also in classroom-based learners. Specifically, as little as three months of L2 immersion has been shown to reduce lexical access in the L1 while enhancing lexical retrieval in the L2 (Linck et al., 2009), and even non-immersive learning contexts can lead to measurable levels of interference in lexical retrieval (Zingaretti et al., 2025) and in the interpretation and processing of overt pronouns in the L1 (Martín-Villena, 2023). Intriguingly, both highly proficient L2 users and L1 attriters have been found to similarly overextend the scope of overt subject pronouns in contexts where null pronouns would be expected (Tsimpli et al., 2004; Sorace & Filiaci, 2006; Belletti et al., 2007), suggesting that there may be convergence between L2 acquisition and L1 attrition. This evidence has led scholars like Sorace (2016) to propose that L1 attrition may be functional to L2 development. While direct investigations of the relationship between the L1 and L2 remain scarce, some findings suggest an inverse relationship in specific linguistic domains such as the lexicon (Linck et al., 2009) and prosody (Zingaretti et al., 2025), while other domains exhibit more complex patterns (Baker, 2024; Zingaretti et al., 2025).

## 1.2. RESEARCH GAP AND AIMS

Taken together, these findings highlight that intensive and meaningful engagement with an L2 can bring about both cognitive and linguistic adaptations, even in late bilinguals. The extent and nature of these changes appear to be modulated by factors such as L2 proficiency, intensity and duration of the learning experience, and sustained use of the L2 over time. Further, linguistic changes appear modulated by the particular domain under investigation. Yet, to date, no study has examined linguistic (in both the L2 and in the L1) and cognitive effects of intensive language learning within the same group of bilinguals. To address this gap, the present study investigated both linguistic and cognitive effects of a one-week, online Italian language course. Specifically, we examined effects in both the L2 (Italian) and the L1 (English) using verbal fluency tasks (Spreen & Strauss, 1998) and focused on attentional inhibition and switching as measured by the subcomponents of the Test of Everyday Attention (Robertson et al., 1994). Crucially, this study did not compare bilinguals to monolinguals. Instead, it focused on within-participant changes across three timepoints (before the course, immediately after and three weeks later), recognising that bilingual performance should not be benchmarked against monolingual norms (Rothman et al., 2023), but rather interpreted in light of bilinguals' own developmental trajectories.

### 1.2.1. Research questions

The questions this pilot study aims to answer are as follows:

RQ1. Are there any improvements in L2 verbal fluency following a one-week language course, and if so, are these inversely correlated with L1 verbal fluency?

RQ2. Are there any cognitive improvements, in terms of attentional inhibition and switching skills, following the course?

RQ3. Do potential changes, either at the linguistic or cognitive level, remain in place three weeks after the end of the course?

### 1.2.2. Hypotheses and predictions

It is important to note that the study was conducted during the COVID-19 pandemic, which necessitated that all instruction and testing be delivered remotely. This online learning environment provided a context for examining whether intensive online language learning can produce effects comparable to those that have been reported in face-to-face language learning environments.

Given the changes in L2 input quality and quantity during the intensive course of Italian, but not in L1 input (as participants remained in an English-dominant environment), we hypothesised that any linguistic effects would be more pronounced in Italian than in English. Further, drawing on previous findings suggesting an inverse relationship between L1 and L2 lexical performance (Linck et al., 2009), we predicted that stronger verbal fluency in L2 Italian might be associated with reduced fluency in L1 English, reflecting a dynamic reconfiguration of the two language systems at the end of the course.

In terms of cognitive outcomes, given that prior studies have shown that even short-term intensive language learning can enhance executive functions (Vega-Mendoza et al., 2015; Bak et al., 2016), we predicted that some cognitive changes would emerge at the end of the course, particularly in the domain of inhibition. In contrast, improvements in attentional switching were considered less likely to occur due to the brevity of the intervention and the online delivery settings.

Finally, given that no ongoing structured practice was implemented after the end of the course, primarily due to practical and logistical constraints during the pandemic period, and unlike in other studies that incorporated post-course engagement (cf. Bak et al., 2016), we hypothesised that any linguistic or cognitive changes would be unlikely to persist at the three-week follow-up testing session.

## 2. METHODOLOGY

### 2.1. PARTICIPANTS

Eighteen undergraduate students enrolled in Italian language programmes at UK universities took part in this study. Participants were recruited in April 2021, in response to a call for a

free one-week online intensive Italian course. The course took place from 12 to 18 April 2021 and was designed to enhance vocabulary and conversational fluency through interactive and immersive online instruction. It was offered during the COVID-19 pandemic, when all teaching had moved to remote formats and UK study-abroad programmes had been halted. Daily instruction included two-hour morning sessions led by L1-Italian language instructors focusing on thematic vocabulary and grammatical structures, followed by one-hour small-group discussions in the afternoon, for a total of approximately 21 hours of live engagement. The programme outline (as well as the full dataset, R code and participant feedback) can be found on the Open Science Framework (OSF) project page (https://osf.io/6bqpx).

Participation in the research study was voluntary and not a requirement for enrolment in the language course. Informed written consent was obtained from all participants, and the study was approved by the School of Philosophy, Psychology and Language Sciences Research Ethics Committee at the University of Edinburgh (reference 236-2021/3), conducted in accordance with the Declaration of Helsinki. Specifically, the study sample included 13 2nd-year and five 3rd-year students. Inclusion criteria required that all participants be L1 speakers of English with Italian as their only actively used L2. Although participants with prior exposure to other languages through formal education were allowed to take part, none of the participants reported using any language other than English or Italian at a conversational level. The mean age of the sample was 19.28 years ($SD = 0.46$). Italian proficiency was assessed at baseline using LexITA, a vocabulary-based measure of L2 proficiency (Amenta et al., 2021). Proficiency scores showed considerable variability (range: 28–78; $M = 47.33$, $SD = 12.98$), which was addressed by including proficiency as a covariate in all subsequent analyses to control for differences between participants.

## 2.2. MATERIALS AND PROCEDURE

### 2.2.1. Verbal fluency tasks

Participants completed a verbal fluency task (VFT) in both English and Italian at each of the three testing sessions. These tasks included phonemic (i.e., 'letter') fluency and semantic (i.e., 'category') fluency, widely used in neuropsychological and bilingualism research (Spreen & Strauss, 1998; Giovannoli et al., 2023). For phonemic fluency, participants were asked to produce as many words as possible beginning with the letters "f", "a" and "s"; for semantic fluency, the categories were animals, food and clothing. Participants had 60 seconds to generate words for each letter or category. The English and Italian tasks were completed on the same day, but scheduled several hours apart (typically, in the morning and evening) in an effort to lessen the likelihood of any immediate cross-linguistic effects that might potentially occur if both tasks were performed back to back. Task order was counterbalanced across participants to reduce potential order effects. Responses were recorded and later coded according to standardised criteria. Namely, words were considered correct if they: fit the given category or began with the correct letter; were real, recognisable words in the target language; included loanwords commonly used and recognised in the language (e.g., *sushi*). Incorrect responses included: intrusions, such as words that did not fit the prompt (e.g., *earrings* for food), non-existent lexical items (e.g. *abdudicate*) or proper nouns for phonemic fluency; repetitions of the same word or morphological variants (e.g., *running* after having said *run*, *sunny* after having said *sun*). Article mismatches in Italian (e.g., *la* [the-FEM] *cane* [dog] instead of *il* [the-MASC] *cane* [dog]) were not penalised, as the instructions had not specified article use.

### 2.2.2. Test of Everyday Attention (TEA)

Attention was assessed using three sub-tests from the Test of Everyday Attention (TEA), a clinical battery developed to assess different facets of auditory attention and widely used in neuropsychological and bilingualism research (Robertson et al., 1994; Vega-Mendoza et al., 2015; Bak et al., 2016). Each sub-test targets a distinct attentional domain: Elevator Task (ET), where participants count tones of identical pitch presented at irregular intervals, targeting auditory sustained attention (seven trials); Elevator Task with Distraction (ET-D), where participants count low-pitched tones while ignoring interspersed high-pitched distractions, targeting auditory selective attention/inhibition (10 trials); and, finally, Elevator Task with Reversal (ET-R), where participants count middle-pitched tones and add or subtract depending on high or low tone cues, targeting auditory attentional switching (10 trials). In keeping with the literature (cf. Vega-Mendoza et al., 2015; Bak et al., 2016) accuracy scores were calculated

for each task separately due to performance possibly differing across tasks, and three alternate versions of each task (A, B and C) were administered across the three testing sessions to minimise practice effects, with all participants receiving the same versions in the same order.

### 2.2.3. Testing

Participants completed three rounds of testing (prior to the course, immediately after the course and three weeks later), allowing participants to serve as their own controls. Testing was conducted remotely via Zoom. While the lack of a separate control group may be viewed as a limitation, this design maximised statistical power in a small sample and is justified by prior research showing that the TEA is resistant to learning effects across repeated testing (Robertson et al., 1994; Bak et al., 2016). Moreover, the observed results on both VFTs and the TEA do not suggest any substantial test–retest effects, as we explain in the Discussion (Section 4).

### 2.3. STATISTICAL ANALYSES

All analyses were conducted in R. Linear mixed-effects models were fitted using the *lme4* package (Bates et al., 2015) with p-values provided by *lmerTest* (Kuznetsova et al., 2017). Post-hoc comparisons were performed using *emmeans* (Lenth, 2025). Data visualisation was carried out using *ggplot2* (Wickham, 2016). Data import and wrangling were performed using *readxl* (Wickham & Bryan, 2025) and *dplyr* (Wickham et al., 2023). Descriptive statistics and correlation analyses used base R functions.

To analyse changes over time, linear mixed-effects models were fitted for each outcome measure: total word production in the L2 Italian and L1 English VFTs, as well as each subtest of the TEA (ET, ET-D and ET-R). All models included timepoint (pre-course, post-course, follow-up) as a fixed effect and participant ID as a random intercept to account for repeated measures. Age and Italian proficiency (LexITA score) were included as covariates in all models. For the VFTs, additional models were run with task type (letter vs. category fluency) and its interaction with timepoint. Model specification was theory-driven: Fixed effects were pre-defined based on study hypotheses, and no stepwise model selection or term removal was applied. Random effects were limited to by-participant intercepts due to sample size and convergence considerations. Model assumptions were checked using residual plots, Q-Q plots and Cook's distance estimates. No violations of normality, homoscedasticity or linearity were observed and no influential outliers were detected. Post-hoc comparisons between timepoints were conducted using Tukey-adjusted pairwise tests.

To assess the relationship between L1 and L2 lexical access, Pearson correlation tests were performed between total word counts in English and Italian VFTs at each timepoint. Assumptions of linearity and bivariate normality were evaluated via Shapiro-Wilk tests and visual inspection of scatterplots with fitted regression lines and 95% confidence intervals. Although one variable (Italian post-course) slightly deviated from normality ($p = 0.047$), the correlations were considered valid given the test's robustness to minor violations.

See OSF project page for full dataset and code (https://osf.io/6bqpx).

# 3. RESULTS

## 3.1. VFT RESULTS

Descriptive statistics for the total number of words produced in the VFTs by timepoint (pre-course, post-course and follow-up) and language (L2 Italian, L1 English) are reported in Table 1, while Table 2 provides a further breakdown by task type, distinguishing between letter and category fluency performance.

| TIMEPOINT | MEAN (SD) | |
|---|---|---|
| | *L2 ITALIAN* | *L1 ENGLISH* |
| **Pre-course** | 53.28 (15.57) | 124.89 (18.24) |
| **Post-course** | 70.44 (16.83) | 124.11 (14.82) |
| **Follow-up** | 70.39 (15.56) | 123.61 (14.75) |

**Table 1** Descriptive statistics (mean and SD) for total words produced in the verbal fluency task by timepoint (pre-course, post-course and follow-up) and language (L2 Italian, L1 English).

| TIMEPOINT | MEAN (SD) | | | |
|---|---|---|---|---|
| | **L2 ITALIAN** | | **L1 ENGLISH** | |
| | *LETTER* | *CATEGORY* | *LETTER* | *CATEGORY* |
| **Pre-course** | 26.00 (9.25) | 27.22 (10.27) | 49.17 (10.73) | 75.72 (11.62) |
| **Post-course** | 31.00 (10.00) | 39.44 (9.56) | 49.06 (9.10) | 75.06 (9.82) |
| **Follow-up** | 32.44 (8.39) | 37.94 (9.50) | 50.06 (9.46) | 73.56 (9.16) |

**Table 2** Descriptive statistics (mean and SD) for words produced in the verbal fluency task by timepoint (pre-course, post-course and follow-up), language (L2 Italian, L1 English) and task type (letter, category).

### 3.1.1 L2 Italian VFT

Performance on the Italian verbal fluency task significantly improved following the course. Relative to the pre-course baseline, total word production increased at both the post-course ($\beta$ = 17.17, $SE$ = 2.07, $t(34)$ = 8.31, $p$ < .001) and follow-up ($\beta$ = 17.11, $SE$ = 2.07, $t(34)$ = 8.28, $p$ < .001) sessions. No significant difference was found between post-course and follow-up ($\beta$ = 0.06, $SE$ = 2.07, $t(34)$ = 0.03, $p$ = .999). LexITA score was a significant positive predictor of verbal fluency ($\beta$ = 0.77, $SE$ = 0.25, $t(15)$ = 3.13, $p$ = .007), whereas age was not ($\beta$ = −6.37, $SE$ = 6.91, $t(15)$ = −0.92, $p$ = .371). See Figure 1 for total word production in L2 Italian over time and Figure 2 for the effect of L2 proficiency on Italian VFT performance.
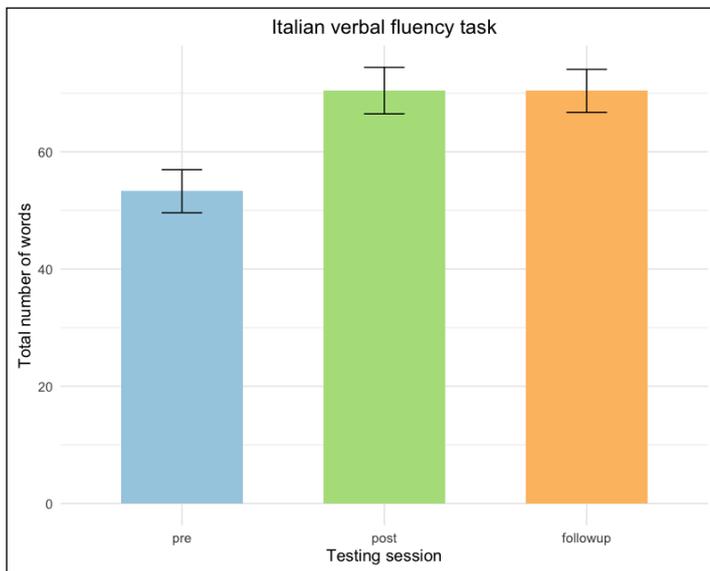


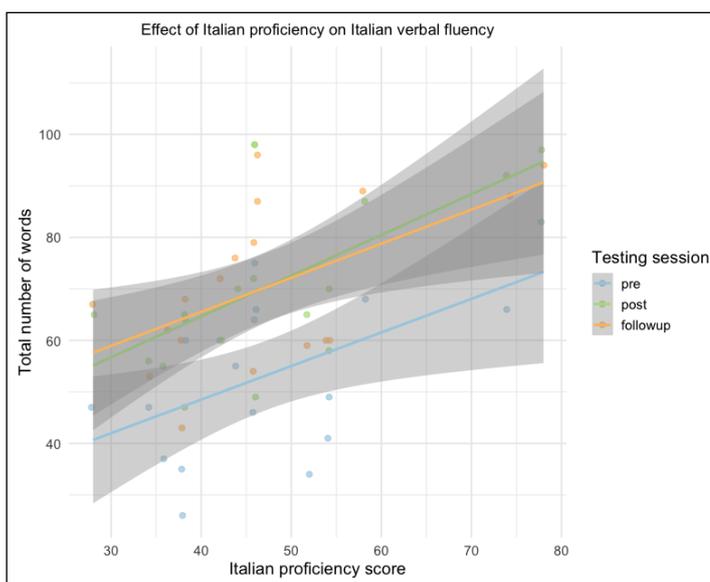**Figure 1** Total word production in L2 Italian verbal fluency over time.



**Figure 2** Effect of Italian proficiency on Italian verbal fluency.

A second model assessed interactions between task type (letter vs. category fluency) and timepoint. There was a significant interaction at post-course ($\beta$ = 7.22, $SE$ = 2.95, $t(85)$ = 2.44, $p$ = .017), reflecting greater improvement in category compared to letter fluency. This

interaction was not significant at follow-up ($\beta$ = 4.28, *SE* = 2.95, *t*(85) = 1.45, *p* = .151). Post-hoc comparisons showed a significant increase in category fluency between pre-course and both post-course ($\beta$ = 12.22, *SE* = 2.09, *t*(85) = 5.85, *p* < .001) and follow-up ($\beta$ = 10.72, *SE* = 2.09, *t*(85) = 5.13, *p* < .001). Improvements in letter fluency were smaller and not statistically significant between pre- and post-course ($\beta$ = 5.00, *SE* = 2.09, *t*(85) = 2.39, *p* = .170) or pre- and follow-up ($\beta$ = 6.44, *SE* = 2.09, *t*(85) = 3.09, *p* = .032). The main effect of LexITA remained significant ($\beta$ = 0.38, *SE* = 0.12, *t*(15) = 3.11, *p* = .007), while age was again non-significant ($\beta$ = −3.17, *SE* = 3.47, *t*(15) = −0.91, *p* = .375). See Figure 3 for the *task type × timepoint* interaction in L2 Italian.
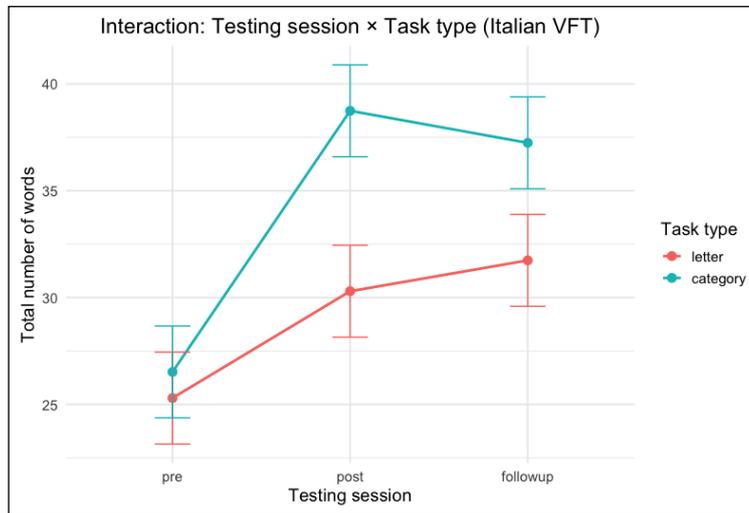


**Figure 3** Interaction between task type (letter vs. category) and timepoint (pre-course, post-course and follow-up) in Italian verbal fluency.

### 3.1.2. L1 English VFT

Performance on the English verbal fluency task remained stable across all sessions. No significant differences in total word production were found between pre-course and post-course ($\beta$ = −0.78, *SE* = 1.66, *t*(34) = −0.47, *p* = .642), pre-course and follow-up ($\beta$ = −1.28, *SE* = 1.66, *t*(34) = −0.77, *p* = .448) or between post-course and follow-up ($\beta$ = −0.50, *SE* = 1.66, *t*(34) = −0.30, *p* = .951). Age was not a significant predictor of performance ($\beta$ = −1.02, *SE* = 8.40, *t*(16) = −0.12, *p* = .905). See Figure 4 for total word production in L1 English over time.
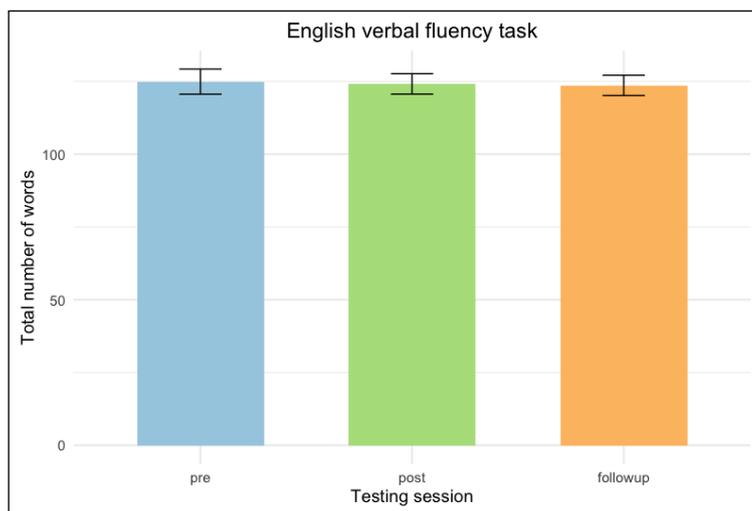


**Figure 4** Total word production in L1 English verbal fluency over time.

A second model assessed interactions between task type and timepoint. There was a robust main effect of task type, with participants consistently producing more words in the category than in the letter fluency condition ($\beta$ = 26.56, *SE* = 2.32, *t*(85) = 11.44, *p* < .001). However, no significant *timepoint × task type* interactions were observed at post-course ($\beta$ = −0.56, *SE* = 3.28, *t*(85) = −0.17, *p* = .866) or follow-up ($\beta$ = −3.06, *SE* = 3.28, *t*(85) = −0.93, *p* = .355), indicating that performance in both task types remained stable over time. Age again did not predict performance ($\beta$ = −0.51, *SE* = 4.20, *t*(16) = −0.12, *p* = .905). See Figure 5 for the *task type × timepoint* interaction in L1 English.

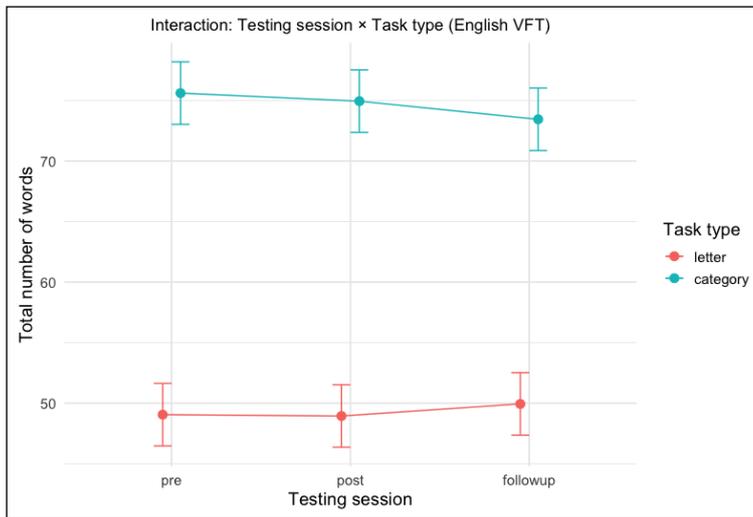**Figure 5** Interaction between task type (letter vs. category) and timepoint (pre-course, post-course and follow-up) in English verbal fluency.

### 3.1.3. Relationship between L2 and L1 verbal fluency

To examine whether performance in L2 verbal fluency might be (inversely) related to verbal fluency performance in the L1, Pearson's correlations were computed between total words produced in the English and Italian verbal fluency tasks at each timepoint. At pre-course, the correlation was not statistically significant ($r = .448$, $p = .062$). A strong positive association emerged at post-course ($r = .791$, $p < .001$), which remained moderate to strong at follow-up ($r = .645$, $p = .004$). Figures 6 and 7 illustrate the significant correlations observed at the post-course and follow-up timepoints, respectively, with fitted regression lines at each timepoint.
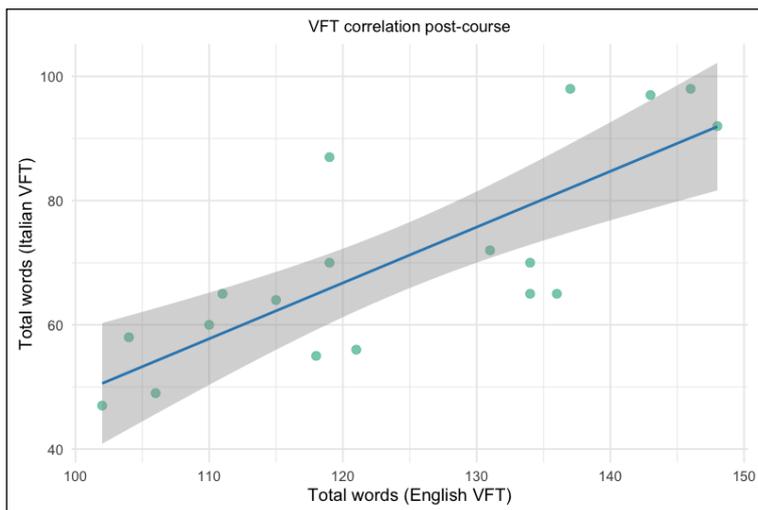


**Figure 6** Correlation between total word production in English and Italian verbal fluency tasks at post-course testing.
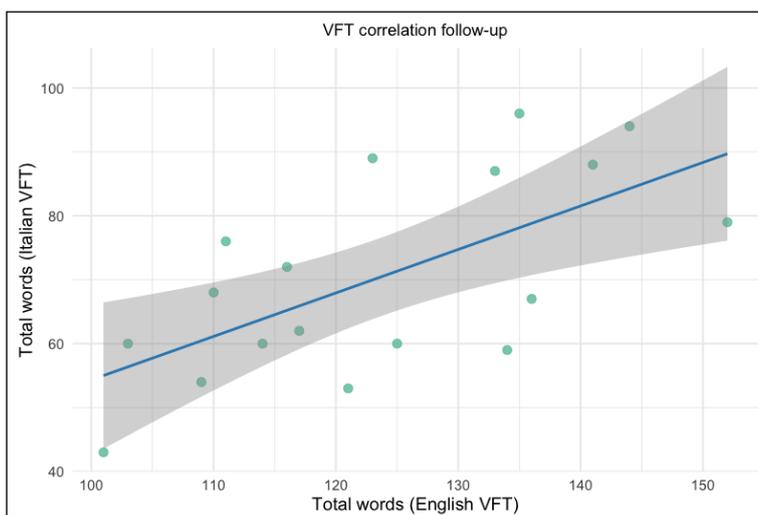


**Figure 7** Correlation between total word production in English and Italian verbal fluency tasks at follow-up testing.

## 3.2. TEA RESULTS

Table 3 summarises mean performance and standard deviations for each TEA sub-test (ET, ET-D, ET-R) across the three testing sessions (pre-course, post-course, and follow-up).

| TIMEPOINT | MEAN (SD) | | |
|---|---|---|---|
| | ELEVATOR TASK | ELEVATOR TASK WITH DISTRACTION | ELEVATOR TASK WITH REVERSAL |
| **Pre-course** | 98.33 (4.85) | 80.56 (21.27) | 80.00 (29.51) |
| **Post-course** | 98.33 (4.85) | 93.89 (10.37) | 88.89 (14.91) |
| **Follow-up** | 96.67 (6.42) | 92.78 (11.27) | 89.44 (11.1) |

### 3.2.1. ET (sustained attention)

Performance in the ET was at near-ceiling levels across sessions (pre: $M = 98.33$, $SD = 4.85$; post: $M = 98.33$, $SD = 4.85$; follow-up: $M = 96.67$, $SD = 6.42$). The mixed-effects model revealed no significant effects of timepoint on accuracy (post vs. pre: $\beta = 0.00$, $SE = 1.68$, $t(34) = 0.00$, $p = 1.000$; follow-up vs. pre: $\beta = -1.67$, $SE = 1.68$, $t(34) = -0.99$, $p = .588$). Neither age ($\beta = -1.15$, $SE = 2.04$, $t(15) = -0.57$, $p = .580$) nor L2 proficiency (LexITA: $\beta = 0.01$, $SE = 0.07$, $t(15) = 0.13$, $p = .897$) significantly predicted performance. These findings indicate stable sustained attention across timepoints, with no effects of the course or individual differences (see Figure 8).
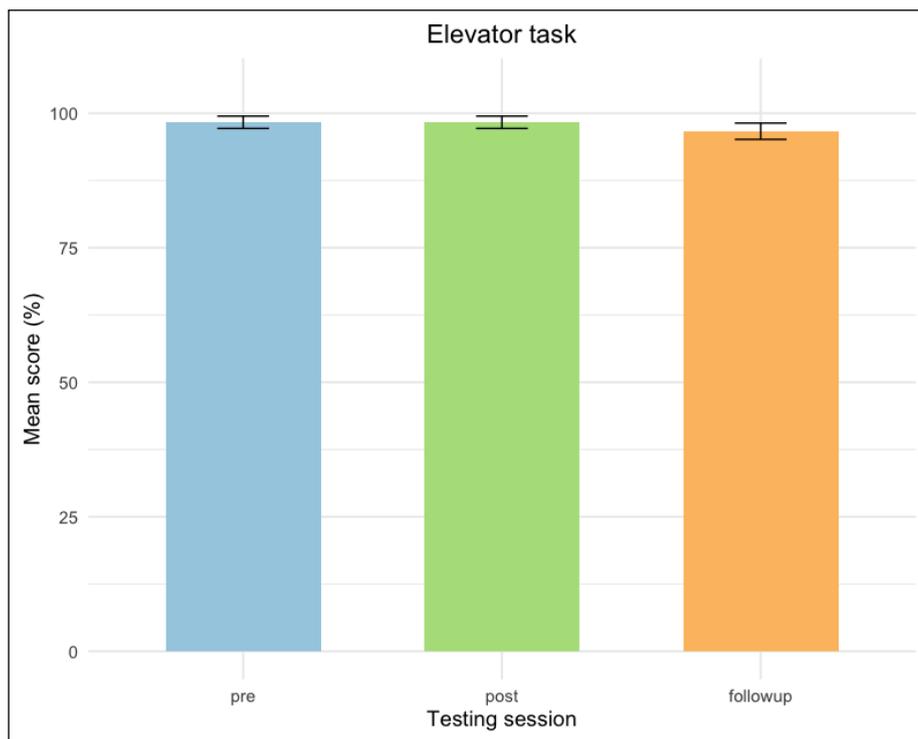
### 3.2.2. ET-D (selective attention/inhibition)

Accuracy improved following the course in the ET-D task (pre: $M = 80.56$, $SD = 21.27$; post: $M = 93.89$, $SD = 10.37$; follow-up: $M = 92.78$, $SD = 11.27$). The model showed a significant effect of timepoint, with both post-course ($\beta = 13.33$, $SE = 3.92$, $t(34) = 3.40$, $p = .005$) and follow-up scores ($\beta = 12.22$, $SE = 3.92$, $t(34) = 3.12$, $p = .010$) significantly higher than pre-course. No difference was observed between post-course and follow-up ($\beta = 1.11$, $SE = 3.92$, $t(34) = 0.28$, $p = .957$). Neither age ($\beta = -10.36$, $SE = 6.30$, $t(15) = -1.65$, $p = .121$) nor LexITA ($\beta = 0.07$, $SE = 0.22$, $t(15) = 0.31$, $p = .760$) were significant predictors. These results indicate improved selective attention and inhibitory control following the course, maintained at follow-up (see Figure 9).
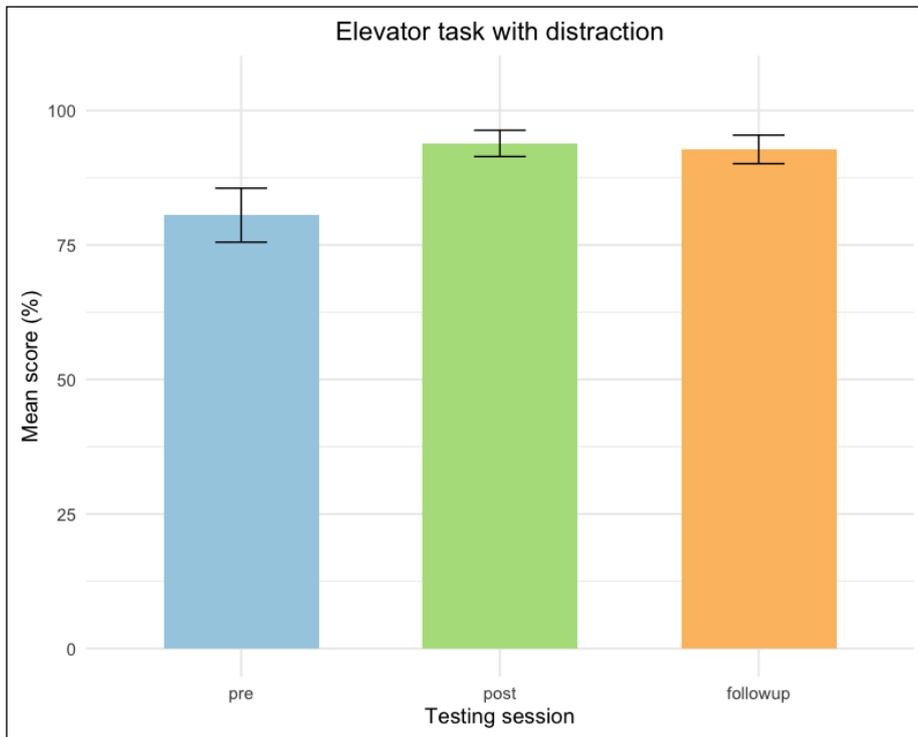
**Figure 9** Mean accuracy score on the Elevator Task with Distraction (measuring auditory selective attention/inhibition) over time.

### 3.2.3. ET-R (attentional switching)

For the ET-R, the numerical increase in accuracy across timepoints (pre: $M$ = 80.00, $SD$ = 29.51; post: $M$ = 88.89, $SD$ = 14.91; follow-up: $M$ = 89.44, $SD$ = 11.10) did not reach statistical significance in the mixed-effects model (post vs. pre: $\beta$ = 8.89, $SE$ = 6.13, $t$(34) = 1.45, $p$ = .327; follow-up vs. pre: $\beta$ = 9.44, $SE$ = 6.13, $t$(34) = 1.54, $p$ = .285). However, individual differences significantly predicted performance: Older age was associated with reduced accuracy ($\beta$ = −18.10, $SE$ = 5.83, t(49) = −3.10, $p$ = .003) and greater L2 proficiency was associated with higher scores ($\beta$ = 0.50, $SE$ = 0.21, $t$(49) = 2.44, $p$ = .019). These results suggest that attentional switching was not improved at the group-level (see Figure 10), but was influenced by individual differences in age and L2 proficiency (see Figures 11 and 12).
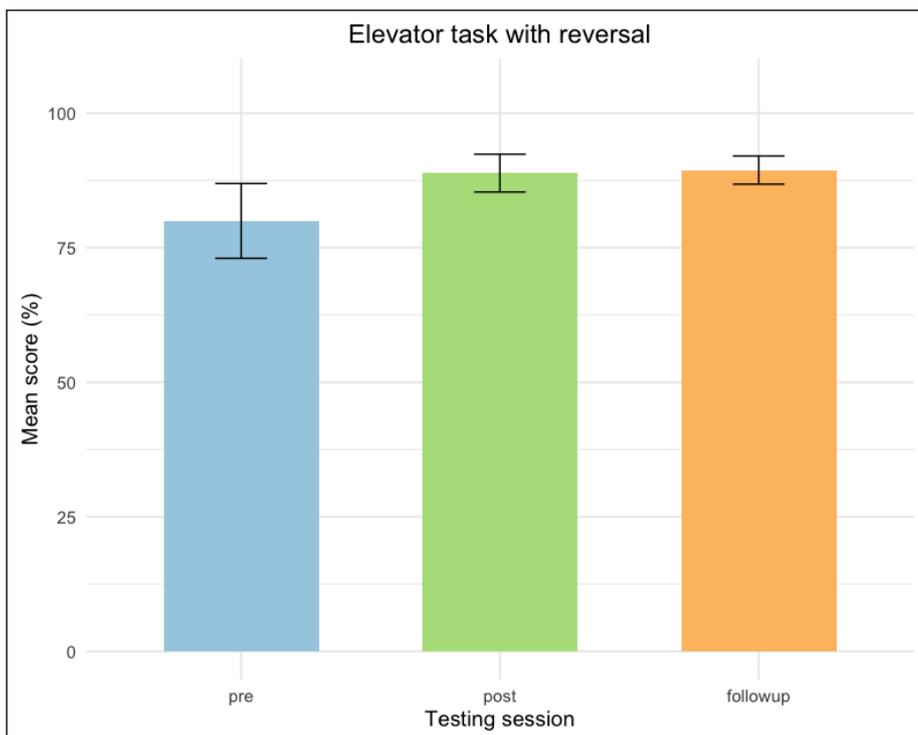


**Figure 10** Mean accuracy score on the Elevator Task with Reversal (measuring auditory attentional switching) over time.
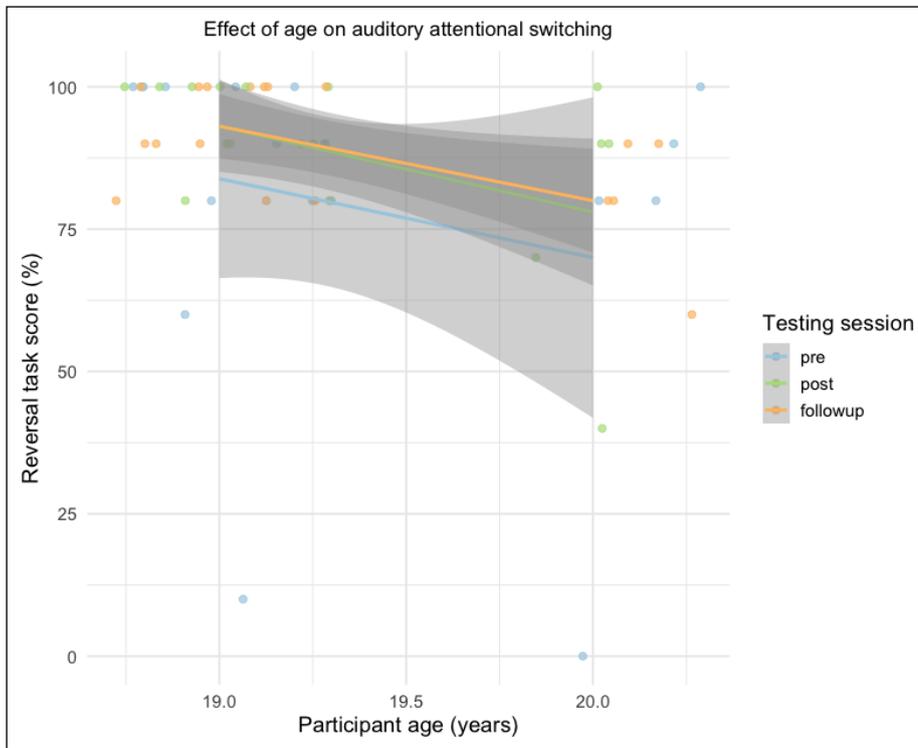
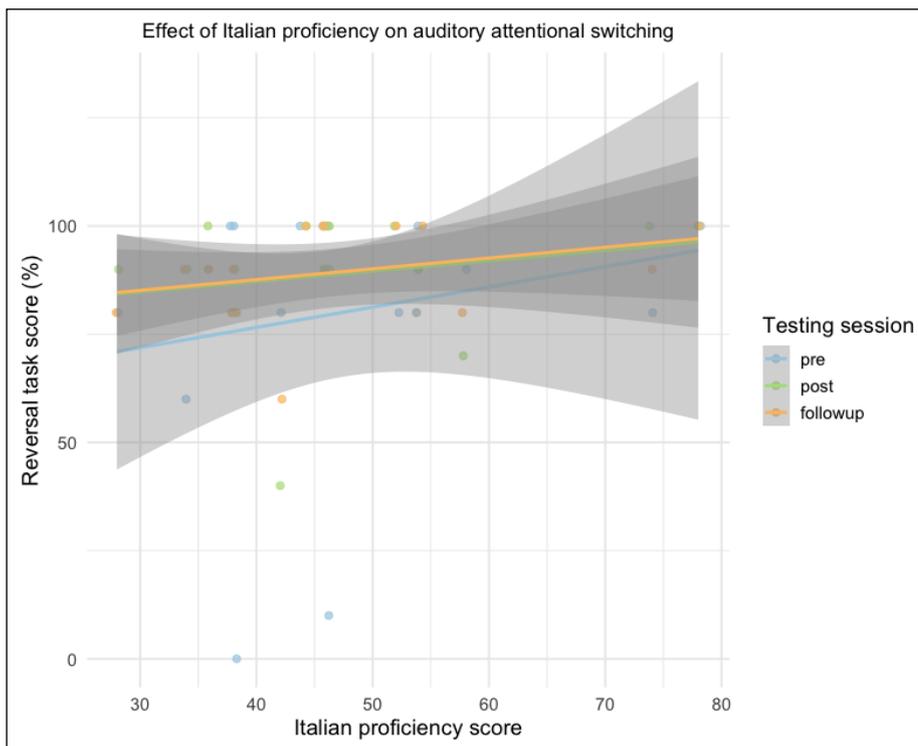**Figure 11** Effect of age on auditory attentional switching.



**Figure 12** Effect of Italian proficiency on auditory attentional switching.

## 4. DISCUSSION

This pilot study examined whether one week of intensive online Italian instruction could yield significant changes in bilinguals' verbal fluency skills in the L2 and the L1, as well as changes in attentional skills (selective attention/inhibition and switching). In response to RQ1, we found significant improvements in L2 verbal fluency that were positively – not inversely – correlated with L1 performance. For RQ2, selective attention/inhibition showed significant improvements while attentional switching did not improve at the group level. Regarding RQ3, both linguistic and cognitive gains persisted at the three-week follow-up. We discuss each of these findings in detail below.

With regards to the first part of RQ1, participants showed a significant increase in the total number of words produced in L2 Italian following the course, with a significant change particularly noticeable for semantic (category) fluency as opposed to phonemic (letter) fluency. Further, higher L2 proficiency predicted better L2 verbal fluency performance. These results suggest that even short-term online L2 learning can strengthen lexical retrieval in an L2, likely by enhancing links between concepts and L2 words through increased L2 exposure and use, which aligns with the predictions made by bilingual lexical access models such as the Revised Hierarchical Model (Kroll & Stewart, 1994; Kroll & Tokowicz, 2005). The observed improvement was most pronounced in semantic fluency, where participants produce words related to a specific category (as we often do in our daily lives), thus revisiting the links between words and concepts (Friesen et al., 2015). This is particularly relevant in the current context, given that the course explicitly targeted vocabulary organised around thematic domains (e.g., food and clothing). In contrast, phonemic fluency, which places greater demands on strategic search and inhibition of irrelevant responses, specifically of related semantic concepts (Luo et al., 2010; Friesen et al., 2015), showed non-significant changes. Taken together, these results suggest that language learning interventions may be especially effective at improving linguistic components of verbal fluency, while aspects that rely more on executive demands may require more sustained or intensive training.

In contrast, verbal fluency in L1 English remained stable across all sessions: Participants consistently performed better in category fluency, but no changes were observed over time, and age was not a significant predictor. Notably, we found no evidence of an inverse L1–L2 relationship in this study (*contra* our predictions): That is, enhanced access to the L2 lexicon did not appear to come at the cost of reduced access to the L1. While prior studies have reported temporary slowdowns in L1 access following L2 immersion (Linck et al., 2009), our participants remained in an L1-dominant environment throughout the course; in this respect, the absence of L1 changes aligns with our hypothesis that short-term, online instruction is less likely to yield reconfigurations of the dominant language system. More unexpectedly, L2 verbal fluency performance was positively associated with L1 verbal fluency performance: Moderate to strong correlations emerged post-course and at follow-up. However, given that L1 attrition has been linked to cognitive load and reduced resources for managing competition from both languages in real time (Sorace, 2016; Zingaretti et al, 2025), the brief, online nature of the course may not have induced sufficient processing demands to trigger such effects. Instead, both languages may have benefited from shared lexical-conceptual resources, particularly in the semantic domain most reinforced during instruction.

Turning to cognitive outcomes, no changes were observed in the Elevator Task, which assessed sustained attention. This was expected, as participants performed at near-ceiling levels across all timepoints, mirroring findings from Bak et al. (2016). In contrast, participants showed significant improvement on the Elevator Task with Distraction (ET-D), which taps selective attention and inhibitory control. This improvement was maintained at follow-up, indicating that even a short, online course can enhance the ability to suppress irrelevant information. One likely explanation is that learners were exposed to increased L2 input and output during the course and had to rely on Italian in real-time interactions while actively suppressing English. Notably, a closer comparison with Bak et al. (2016) reveals that participants in the current study began with a lower baseline in ET-D ($M = 80.56$) than those in Bak et al. ($M = 88.79$), suggesting they had more room to improve.

By contrast, no significant change was observed in the Elevator Task with Reversal (ET-R), which requires switching between mental sets based on auditory cues. This diverges from Bak et al. (2016), where ET-R showed the largest post-course gains (from $M = 59.7$ to $M = 78.5$). In the present study, however, baseline ET-R performance was already relatively high ($M = 80.0$), possibly limiting the potential for measurable change. Moreover, task switching and inhibition rely on partially distinct executive processes with different demands (Miyake et al., 2000), and switching often incurs great performance costs even in high-functioning individuals, due to the need for flexible, moment-to-moment adjustments in cognitive control (Monsell, 2003). Moreover, other studies found enhanced switching only in more advanced learners, such as fourth-year university language students when compared to those in their first year (Vega-Mendoza, et al., 2015). As such, it seems plausible to suggest that improvements in switching may require longer or more immersive engagement with the L2 than what was provided in

the current study. Nonetheless, it is interesting to note that switching performance could be predicted by individual differences: Higher L2 proficiency was associated with better accuracy, while older age was associated with poorer performance. This suggests that although switching did not improve overall at the end of the course, it remains closely linked to individual differences in age and language proficiency.

Importantly, the observed gains in both L2 verbal fluency and selective attention were not transient, but remained stable three weeks after the course ended. This mirrors findings from Bak et al. (2016), who reported lasting cognitive improvements following a similarly brief language intervention. Although the present design did not include a passive or active control group, the selective nature of the effects emerging only in certain subcomponents (e.g., category but not letter fluency; ET-D but not ET-R) argues against a simple regression to the mean. Future research should address these limitations by recruiting larger and more diverse samples, including control groups, and expanding the scope of investigation to include additional linguistic domains (e.g., syntax, phonology) and a broader range of cognitive functions (e.g., working memory).

## 5. CONCLUSION

The findings of this pilot study provide preliminary evidence that even short-term, online L2 learning can confer benefits to both language and cognition. Specifically, a one-week online Italian course was enough to yield improvements in L2 verbal fluency, especially semantic fluency, with no changes in L1 verbal fluency; significant improvements in selective attention/ inhibition were also reported, with linguistic and cognitive changes persisting at follow-up, three weeks after the end of the course. Our findings contribute to ongoing debates on the scope and timescale of language-related cognitive adaptations by showing that measurable benefits can occur outside immersive contexts and over a short online intervention, extending prior work on intensive language learning and executive functions (e.g., Bak et al., 2016; Vega-Mendoza et al., 2015) and on L1–L2 interactions within individuals (e.g., Sorace, 2016; Zingaretti et al., 2025).

Interestingly, these quantitative insights are complemented by further insight from participant feedback (see OSF project page: https://osf.io/6bqpx). 82% of respondents reported perceived gains in their Italian, particularly in vocabulary and fluency. Three participants also commented on changes in the use of both languages, such as feeling "compelled to think more in Italian", "thinking of Italian words instead of English when writing, speaking and thinking in English" and "beginning to identify things … less in English and more in Italian, during and after the course". These introspective reports highlight the experiential dimension of short-term, intensive L2 learning and call for further research into its potential to dynamically influence both language(s) and cognition.

Taken together, these findings support the feasibility of short, intensive online L2 learning and motivate larger, controlled and preregistered studies to probe the causality, durability and scope of the effects of intensive (online) language learning on language and cognition found in the present pilot study.

## AUTHOR AFFILIATIONS

**Mattia Zingaretti** ⓘ orcid.org/0000-0002-5565-2538

Department of Linguistics and English Language, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, Scotland; York St John University, York, United Kingdom

**Roberta Spelorzi** ⓘ orcid.org/0000-0002-4882-4426

Department of Linguistics and English Language, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom

**Taranjit K. Sohanpaul** ⓘ orcid.org/0009-0002-1098-6822

Department of Languages and Linguistics, School of Education, Language and Psychology, York St John University, York, United Kingdom

**Antonella Sorace** ⓘ orcid.org/0000-0002-8812-1979

Department of Linguistics and English Language, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom

## REFERENCES

**Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., Jaydip, R. C.,** & **Subhash, K.** (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology, 81*(22), 1938–1944. https://doi.org/10.1212/01.wnl.0000436620.33155.a4

**Amenta, S., Badan, L.,** & **Brysbaert, M.** (2021). LexITA: A quick and reliable assessment tool for Italian L2 receptive vocabulary size. *Applied Linguistics, 42*(2), 292–314. https://doi.org/10.1093/applin/amaa020

**Bak, T. H., Long, M. R., Vega-Mendoza, M.,** & **Sorace, A.** (2016). Novelty, challenge, and practice: The impact of intensive language learning on attentional functions. *PLOS ONE, 11*(4), e0153485. https://doi.org/10.1371/journal.pone.0153485

**Baker, L.** (2024). *The role of linguistic input in adult grammars: Modelling L1 morphosyntactic attrition* [Unpublished doctoral thesis]. University of Southampton.

**Bates, D., Mächler, M., Bolker, B.,** & **Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

**Belletti, A., Bennati, E.,** & **Sorace, A.** (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from nearnative Italian. *Natural Language & Linguistic Theory, 25*, 657–689. https://doi.org/10.1007/s11049-006-9012-7

**Bialystok, E.** (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*(3), 636–644. https://doi.org/10.1111/1467-8624.00046

**Bialystok, E., Craik, F. I. M.,** & **Freedman, M.** (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459–464. https://doi.org/10.1016/j.neuropsychologia.2006.10.010

**Costa, A., Hernández, M.,** & **Sebastián-Gallés, N.** (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*(1), 59–86. https://doi.org/10.1016/j.cognition.2006.12.013

**de Bruin, A., Treccani, B.,** & **Della Sala, S.** (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science, 26*(1), 99–107. https://doi.org/10.1177/0956797614557866

**Friesen, D. C., Luo, L., Luk, G.,** & **Bialystok, E.** (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience, 30*(3), 238–250. https://doi.org/10.1080/23273798.2014.918630

**Giovannoli, J., Martella, D.,** & **Casagrande, M.** (2023). Executive functioning during verbal fluency tasks in bilinguals: A systematic review. *International Journal of Language & Communication Disorders, 58*(4), 1316–1334. https://doi.org/10.1111/1460-6984.12855

**Gollan, T. H., Montoya, R. I.,** & **Werner, G. A.** (2002). Semantic and letter fluency in SpanishEnglish bilinguals. *Neuropsychology, 16*(4), 562–576. https://doi.org/10.1037/0894-4105.16.4.562

**Gollan, T. H., Montoya, R. I., Fennema-Notestine, C.,** & **Morris, S. K.** (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220–1234. https://doi.org/10.3758/BF03193224

**Green, D. W.** (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition, 1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

**Grosjean, F.** (2013). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (pp. 5–25). Wiley-Blackwell.

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastián-Gallés, N. (2013). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition, 16*(1), 275–287. https://doi.org/10.1017/S1366728912000467

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127*(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging, 23*(1), 70–78. https://doi.org/10.1037/0882-7974.23.1.70

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing: Looking back and to the future. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 531–553). Oxford University Press. https://doi.org/10.1093/oso/9780195151770.003.0030

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2025). *emmeans: Estimated marginal means, aka least-squares means* (Version 1.11.2) [R package]. https://doi.org/10.32614/CRAN.package.emmeans

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in secondlanguage learning. *Psychological Science, 20*(12), 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*(1), 29–41. https://doi.org/10.1016/j.cognition.2009.08.014

Martín-Villena, F. (2023). *L1 Morphosyntactic Attrition at the Early Stages: Evidence from Production, Interpretation, and Processing of Subject Referring Expressions in L1 Spanish L2 English Instructed and Immersed Bilinguals* (Doctoral thesis, Universidad de Granada). https://hdl.handle.net/10481/81920

Michael, E. B., & Gollan, T. H. (2005). Being and becoming bilingual. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). Oxford University Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The Test of Everyday Attention*. Thames Valley Test Company.

Rothman, J., Bayram, F., DeLuca, V., Di Pisa, G., Duñabeitia, J. A., Gharibi, K., Hao, J., Kolb, N., Kubota, M., & Kupisch, T. (2023). Monolingual comparative normativity in bilingualism research is out of "control": Arguments and alternatives. *Applied Psycholinguistics, 44*(3), 316–329. https://doi.org/10.1017/S0142716422000315

Schmid, M. S., & Köpke, B. (2017a). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism, 7*(6), 637–642. https://doi.org/10.1075/lab.17058.sch

Schmid, M. S., & Köpke, B. (2017b). When is a bilingual an attriter? Response to the commentaries. *Linguistic Approaches to Bilingualism, 7*(6), 721–755. https://doi.org/10.1075/lab.17059.sch

Schmid, M. S., & Köpke, B. (Eds.). (2019). *The Oxford handbook of language attrition*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198793595.001.0001

Sorace, A. (2016). Referring expressions and executive functions in bilingualism. *Linguistic Approaches to Bilingualism, 6*(5), 669–684. https://doi.org/10.1075/lab.14018.sor

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in nearnative speakers of Italian. *Second Language Research, 22*(3), 339–368. https://doi.org/10.1191/0267658306sr271oa

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). Oxford University Press.

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition. *Frontiers in Psychology, 2*, 123. https://doi.org/10.3389/fpsyg.2011.00123

Tsimpli, I. M., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism, 8*(3), 257–277. https://doi.org/10.1177/13670069040080030601

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, nonbalanced bilingualism on cognitive performance. *Cognition, 137*, 40–46. https://doi.org/10.1016/j.cognition.2014.12.008

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. https://doi.org/10.1007/978-3-319-24277-4_9

Wickham, H., & Bryan, J. (2025). *readxl: Read Excel files* (Version 1.4.5) [R package]. https://doi.org/10.32614/CRAN.package.readxl

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation* (Version 1.1.4) [R package]. https://doi.org/10.32614/CRAN.package.dplyr

Zingaretti, M., Chondrogianni, V., Ladd, D. R., & Sorace, A. (2025). L1 attrition vis-à-vis L2 acquisition: Lexicon, syntax-pragmatics interface, and prosody in L1-English L2-Italian late bilinguals. *Languages, 10*(9), 224. https://doi.org/10.3390/languages10090224

WHITE ROSE
UNIVERSITY PRESS
Universities of Leeds, Sheffield & York