Bilingualism: Language and Cognition

cambridge.org/bil

Verbal fluency in Greek: Performance differences between L1Greek-L2English late bilingual and Greek monolingual speakers

Dimitra Lazaridou-Chatzigoga[1,2] and Artemis Alexiadou[3,4]

[1]School of Politics, Philosophy and Area Studies & School of Psychology, University of East Anglia, Norwich, UK; [2]Department of Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK; [3]Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin, Germany and [4]Institute of German Language and Linguistics, Humboldt University of Berlin, Berlin, Germany

Research Article

Cite this article: Lazaridou-Chatzigoga, D. and Alexiadou, A. (2024). Verbal fluency in Greek: Performance differences between L1Greek-L2English late bilingual and Greek monolingual speakers. *Bilingualism: Language and Cognition*, 1–8 https://doi.org/10.1017/S1366728924000671

Received: 20 September 2022
Revised: 14 August 2024
Accepted: 18 August 2024

Keywords:
verbal fluency; bilingualism; language attrition; Greek/English; bilinguals; formal fluency; semantic fluency

Corresponding author:
Dimitra Lazaridou-Chatzigoga;
Email: dl518@cam.ac.uk

This article has earned badge for transparent research practices: Open Data Badge. For details see the Data Availability Statement.

Abstract

Verbal fluency data for semantic (animals, fruit and vegetables and objects) and formal fluency (*Χ* [Chi], *Σ* [Sigma] *Α* [Alpha]) were collected from 32 L1Greek-L2English late bilingual speakers and 32 Greek monolinguals. The verbal fluency task has been used in both language attrition and bilingualism studies. Language attrition studies, which mostly employ only the semantic task, show that bilinguals perform worse than monolinguals. In bilingualism studies, which employ both the semantic and formal tasks, we find greater variance and the results are mixed (bilinguals perform similarly, better or poorly compared to monolinguals). In our study, we investigated quantitative measures (number of correct responses) and strategic processes (clustering, switching). In the quantitative measures, monolinguals outperformed bilinguals in both tasks with the difference being more pronounced in the semantic task. In clustering, both groups behaved similarly, while in switching monolinguals performed better than bilinguals. The implications of these results are discussed.

Highlights

- We collected verbal fluency data from Greek monolinguals and bilinguals (L1Greek/L2 English).
- Verbal fluency included semantic (three categories) and formal fluency (three letters).
- Greek monolinguals outperformed bilinguals in number of correct responses.
- Both groups behaved similarly in clustering.
- Monolinguals performed better than bilinguals in switching.

1. Introduction

1.1. General introduction

Verbal fluency tasks are extensively used by clinical neuropsychologists to assess lexical access difficulties/word retrieval efficiency in order to screen for impairment in language ability and executive control functioning in conditions such as aphasia, dementia and schizophrenia among others (e.g., Spreen & Benton, 1977). More recently, verbal fluency tasks have been used by scholars studying language attrition (see discussion in Schmid, 2011; Schmid & Jarvis, 2014) and bilingualism (see Bialystok et al., 2008; Patra et al., 2020 inter alia). The verbal fluency task measures the ability to generate as many words as possible in a fixed time provided (usually 60 s) based on a given criterion. Two types of criteria are usually used: the semantic task measures the ability to retrieve and generate words on the basis of a semantic category (e.g., animals, fruit), while the formal (or letter/phonemic) task measures the ability to retrieve and generate words based on the initial letter of the word (e.g., words that begin with F; most commonly F, A and S for English). Both tasks require a combination of language abilities and executive control abilities. Performance on a verbal fluency task reflects the integrity of the mental lexicon (i.e., the system that contains information regarding a word's meaning, pronunciation and other characteristics as well as word associations), the efficiency of retrieval strategies, as well as of self-monitoring and inhibition of inappropriate responses. The search criteria in semantic fluency resemble the way words are generated in everyday activities. For instance, if participants are asked to generate items belonging to the semantic category of fruit, they might try to think of the items in their fridge or at their local greengrocer. Thus, the demands for semantic fluency are similar to the way semantic memory is structured as concepts are clustered along semantic properties (Friesen et al., 2015; Luo et al., 2010). On the contrary, the search criteria for formal fluency, i.e., to generate words that begin with a certain letter, do not resemble practices in everyday activities and are not similar to

CAMBRIDGE UNIVERSITY PRESS

the structure of semantic memory, as lexical entries are not listed alphabetically. Furthermore, in the formal fluency task, participants must employ additional inhibition in order to suppress semantically related words and they have to come up with novel strategies in order to generate words (Friesen et al., 2015; Luo et al., 2010). In terms of the different demands that the two tasks require, semantic fluency is associated more with language abilities such as vocabulary size and the structure of the semantic lexical network, while formal fluency is associated more with executive control abilities such as working memory, response inhibition and self-monitoring (see discussion in Luo et al., 2010 inter alia). More recent work argues that while semantic and formal fluencies correlate in a similar way across measures of executive functioning (working memory, fluid reasoning and shifting/updating), the results show that a larger executive component is involved in the semantic task (Aita et al., 2019). Thus, the discussion of the relation of semantic and formal fluencies to language and executive functioning is not settled yet.

In language attrition research, the verbal fluency task is a popular tool due to its simple administration across languages and its reliability across different populations (Roberts & Le Dorze, 1998). It has been furthermore characterised as a rewarding tool since it mostly produces significant findings in a field, where findings are often not clear-cut. In the standard version of the task, the participant's score is the number of correct responses (i.e., the number of correct responses produced in 60 s excluding errors). In language attrition studies, the most common format of the verbal fluency task has been the semantic task. Usually, two semantic categories are used, 'animals' and 'fruit and vegetables'. In a number of studies, attriters (or late sequential bilinguals) have been found to show lower scores in the semantic task than control groups (or monolinguals) when their L1 is tested (Cherciov, 2011; Keijzer, 2007; Schmid, 2007; Waas, 1996; Yağmur, 1997). Cherciov (2011) tested bilinguals in both languages and they performed better in L1-Romanian than in L2-English. Employing only the semantic task might under-represent attriters' performance as the semantic task relies more on linguistic abilities that might have been altered because of their bilingual status[1]. Discussion in the literature suggests that a formal task may give rise to interesting results in an attritional setting (Schmid, 2011), as formal fluency also gives rise to more variation within normal populations as well (Roberts & Le Dorze, 1998). There are a few studies that have used both semantic and formal tasks in both languages with interesting findings. Ammerlaan (1996) found that Dutch–English bilinguals performed better in L2-English than in L1-Dutch in both tasks and Opitz (2011) found that German–English bilinguals produced fewer words than monolinguals in both L1-German and L2-English and in both tasks, but the differences did not reach statistical significance, which might be attributed to the small sample size (13 attriters, 17 controls). More recently, Shishkin and Ecke (2018) investigated both semantic and formal fluencies in two groups of Russian-English bilinguals and found no substantial attrition in L1 fluency. Their design did not include though a monolingual group; thus, it is not possible to know how their performance differs from monolingual Russian speakers. More research employing both tasks in bilinguals and the inclusion of a monolingual group is clearly needed to see the respective contribution of each component of verbal fluency.

In bilingualism research, the verbal fluency task has been mainly used to inform the debate of linguistic and executive control differences between monolingual and bilingual speakers, and it has yielded mixed results. With respect to semantic fluency, studies typically show that monolinguals outperform bilinguals in the latter's L2 as the former generate more items (Gollan et al., 2002; Kormi-Nouri et al., 2012; Paap et al., 2017; Rosselli et al., 2000; Sandoval et al., 2010). This difference between monolinguals and bilinguals disappears though when the two groups are matched on (receptive) vocabulary (Bialystok et al., 2008; Luo et al., 2010; Paap et al., 2017; Patra et al., 2020). Results on formal fluency are more varied. There have been claims in the literature that monolinguals outperform bilinguals (Sandoval et al., 2010), that the two groups do not differ (Bialystok et al., 2008; Kormi-Nouri et al., 2012; Paap et al., 2017; Portocarrero et al., 2007; Vega-Mendoza et al., 2015 among others) and that bilinguals outperform monolinguals (Friesen et al., 2015; Kormi-Nouri et al., 2012; Ljungberg et al., 2013; Patra et al., 2020).

It is important to highlight here that whereas the majority of the language attrition studies reviewed above focus on the L1, the majority of the bilingualism studies focus on the L2. This might explain the greater homogeneity we observe in the language attrition studies, which also mostly employ only the semantic task, as opposed to the greater variance found in the bilingualism studies that usually employ both the semantic and the formal tasks.

Apart from the number of correct responses, some studies have employed additional quantitative measures such as mean subsequent-response latencies (i.e., measuring the time between the first response and the onset of each subsequent response, with the mean subsequent-response latency indicating the point at which half of the responses have been produced – also called "fulcrum" in Sandoval et al., 2010) and time-course analyses (i.e., measuring different variables on the timing of the responses) in Friesen et al. (2015) and Luo et al. (2010) or average productivity (i.e., each task is divided into six segments of equal length and the average productivity in each segment is calculated) in Schmid and Jarvis (2014). The use of finer-grained information on semantic verbal fluency such as error types, speech breaks and semantic relatedness has also been recently investigated (Amunts et al., 2021).

Qualitative measures or strategic processes, as we will refer to them in this article, have also been employed in order to investigate the cognitive strategies used to complete the task successfully. Analyses of the processes of generating words have shown that words are produced in spurts or temporal clusters rather than at a consistent rate throughout the duration of the task (Gruenewald & Lockhead, 1980). Successful production of words seems to rely on identifying words based on semantic relations on the semantic fluency task (e.g., first naming pets, then switching to farm animals, then to animals of the jungle) or on phonological relations on the formal fluency task (e.g., words that start with the same two letters, then switch to words that share the same ending, then to words that rhyme). The process of organising words into semantically or phonologically related subcategories has been characterised as *clustering* (Robert et al., 1997; Troyer et al., 1997). A related process is referred to as *switching*, which describes the ability to shift efficiently from one subcategory to another. Robert et al. (1997) and Troyer et al. (1997) found that both clustering and switching strategies are equally highly correlated with the total number of words produced on semantic fluency, indicating that both are

[1]In addition, as we will see, based on the results of our study, using only semantic categories such as 'animals' and 'fruit and vegetables' might be under-representing the bilinguals' semantic verbal fluency as these two categories produce different results from the 'objects' category.

https://doi.org/10.1017/S1366728924000671 Published online by Cambridge University Press

important components for optimal performance on this task. In contrast, switching is more highly correlated than clustering with the total number of words produced in formal fluency, indicating that switching is more important for optimal performance on this task. Filippetti and Allegri (2011) – with data from Spanish-speaking children – on the other hand, argue that clustering accounts for more variance in semantic fluency, whereas switching accounts for more variance in formal fluency.

There is scarcely any literature comparing bilinguals to monolinguals on the strategic processes of clustering and switching, with the notable exception of Patra et al. (2020), who investigated the L2 (English) of Bengali–English bilinguals. They hypothesised that vocabulary-matched bilinguals would produce equal cluster size and a larger number of switches. Their results showed that for both monolinguals and bilinguals, cluster size on the semantic task was bigger than cluster size on the formal task. More importantly, bilinguals produced significantly larger clusters than the monolinguals on the formal task but were comparable in cluster size on the semantic task. Their results did not confirm all their predictions, which they attribute to a strategy that bilinguals developed to compensate for the more demanding formal task. There were no differences in switching between monolinguals and bilinguals, which they found surprising and thus they suggest that perhaps this is not a strategy used by bilinguals. More research is clearly needed. Both strategies can further inform our understanding of the similarities and differences of the performance in bilinguals versus monolinguals.

As far as we know, there is no other study reporting data from a verbal fluency task (VFT) in Greek bilinguals. There is one study though that assessed 300 monolingual adults, providing the first normative data for the Greek population (Kosmidis et al., 2004). On the semantic task, they used the following three semantic categories: animals, fruit and objects. On the formal task, they used the following three Greek letters: *Χ* (Chi), *Σ* (Sigma) and *Α* (Alpha). The selection of the letters was based on the ratio of words in Greek that start with these three letters relative to the total number of words in a Greek dictionary, in correspondence with the ratio of words in English that start with the letters *F*, *A* and *S*, relative to the total number of words in an English dictionary (the *FAS* task was first implemented in English, see, e.g., Benton, 1968). Demographic variables such as level of education, age and to some extent, sex contributed to verbal fluency, in line with previous literature on English.

For the present study, we decided to use the categories of Kosmidis et al. (2004) and test both semantic and formal fluencies as we established that more research is needed in order to measure their respective contribution in verbal fluency and to see how potential attriters would perform in each of the tasks compared to monolinguals. For reasons of time, we focused only on the L1 of the bilingual speakers. We used both quantitative measures and strategic processes in order to get a fuller picture of lexical access in bilinguals.

1.2. The current study

We compared performance in the verbal fluency task in two groups of healthy adult participants: 32 L1-Greek L2-English late bilinguals and 32 Greek monolinguals. We collected data on semantic (animals, fruit and vegetables, objects) and formal (*Χ* [Chi], *Σ* [Sigma] and *Α* [Alpha]) fluency by giving each participant 60 s for each category or letter in Greek. For the bilinguals, we collected information on relevant variables for their bilingualism, such as language history, language use, language proficiency and language attitudes for each language. Based on this, language dominance was computed.

We formulated our hypotheses based on the literature review above. We predicted that bilinguals would show lower scores than monolinguals in semantic fluency in line with the majority of language attrition and bilingualism studies. The picture for formal fluency is more undecided as every possible view has been defended and supported by evidence in the bilingualism studies. As highlighted above, these results rest on the bilinguals' L2. Only Opitz (2011), as far as we know, has tested whether this would also be the case for the bilinguals' L1 with results that show that bilinguals received lower scores than monolinguals. Their sample size was small though and the differences did not reach statistical significance. Following a cautionary line, we predicted that bilinguals in our study would perform equally or worse than monolinguals on formal fluency. In line with our predictions regarding the number of correct responses, we were expecting that bilinguals would show a smaller cluster size in the semantic task and potentially equal cluster size in the formal task than monolinguals. With respect to switching, we were expecting that bilinguals would perform similarly to monolinguals in number of switches in both semantic and formal fluencies.

2. The study

2.1. Methods

2.1.1. Participants

We assessed 32 L1-Greek L2-English late bilingual speakers (20 females, 12 males; aged 27–46; mean age 37.34; SD = 5.10) and 32 L1-Greek monolingual speakers (21 females, 11 males; aged 22–44; mean age 31.125; SD = 6.04). All participants were highly educated (they had at least a university degree). They were all recruited via online groups and word-of-mouth and gave their written informed consent to participate in the study following the ethics protocol of the Ethics Committee of the German Linguistic Society (Deutsche Gesellschaft für Sprachwissenschaft). Participants received monetary compensation for their participation.

The bilingual participants resided in London in the United Kingdom, having migrated to the United Kingdom from Greece as adults to study and/or work. The monolingual participants resided in Athens or Thessaloniki in Greece. The bilinguals had 6–22 years of residence in the United Kingdom (mean = 13.13; SD = 5.53) and spoke both Greek and English fluently. The monolinguals used only Greek in their day-to-day life and considered themselves monolinguals even though they had some knowledge of other languages due to school education. Their scores on a close test measuring proficiency in Greek (based on Tsimpli, p.c.) were similar: bilinguals scored 43.44 (SD 3.82), while monolinguals scored 45.4 (SD 2.78). Because of the constraints of recruiting, the groups were not matched on age and proficiency in Greek, but they were as close as possible. We included both as predictors in our models.

2.1.2. Measures of bilingualism

The bilingual participants completed the Bilingual Language Profile (BLP) questionnaire (Birdsong et al., 2012), where they self-assessed language history, current use, proficiency and attitudes for both languages. All participants started learning Greek from birth at home and English in their childhood at school or with private tutors (mean = 8.16, SD = 1.68, range = 4–13). The

https://doi.org/10.1017/S1366728924000671 Published online by Cambridge University Press

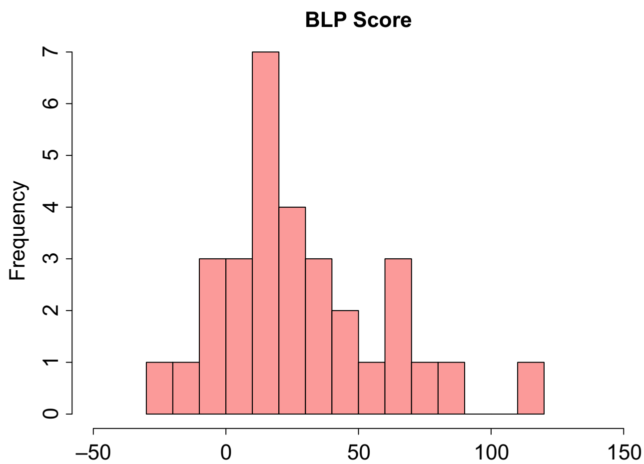

Figure 1. BLP scores across participants.

participants obtained a mean BLP language dominance score of 28.44 (SD = 31.33, range: −22.25 to 117.15) calculated based on weighted scores of each component of the questionnaire for each language. Figure 1 shows the distribution of BLP scores, which reveals that Greek-dominant participants (BLP > 0) outnumber English-dominant ones. The majority of the individuals in the sample could be characterised as balanced bilinguals (as most scores are close to 0) with a tendency towards dominance in Greek.

2.1.3. Procedure

The study was administered by a Greek native speaker and was conducted in Greek. It included a verbal fluency task that was comprised by a semantic task and a formal task. Both tasks were part of a 2 h 1-to-1 session with the researcher. On the semantic task, we asked participants to generate as many different words as possible belonging to each of the following three semantic categories: animals, fruit and vegetables, and objects. Participants were told to avoid repetitions. On the formal task, we asked participants to generate as many different words as possible beginning with each of the following three Greek letters: *Χ* (Chi), *Σ* (Sigma) and *Α* (Alpha) (following Kosmidis et al., 2004). On the formal task, participants were additionally instructed to avoid proper nouns and variations of the same word. We illustrated this by providing the following words to be avoided for the letter E (Epsilon): *Eladha* 'Greece', *Eleni* 'Eleni' and variations of the same word (e.g., words from the same stem, such as *epiloghi* 'selection', *epilektos* 'selected' after *epilegho* 'select'). Participants had 60 s for each trial. Both tasks were recorded on a computer via Audacity. No guidelines were given to the participants on how to organise their word search and production. The semantic task was administered prior to the formal task, and categories and letters were administered in the above-mentioned order for all participants, following Kosmidis et al. (2004).[2]

[2]In this, we followed Kosmidis et al. (2004), as they represent the norm for Greek, and we wanted our results to be comparable to theirs. They used the same order for all participants (first the semantic task and then the formal fluency task). As a reviewer points out, this might have affected the results, which, as we will see, show that participants across the two groups produced less words overall in the formal task than in the semantic task, possibly due to fatigue effects. In the future, it would be good to flip the order for half of the participants to be able to exclude that possibility.

2.1.4. Data scoring and analysis

Responses were transcribed and then scored by one of the researchers and a research assistant. The intraclass correlation coefficient was 0.995, which indicates excellent reliability. Good absolute agreement between the raters was observed while using the two-way random-effect models and a single rater with a p-value < .001. A given item would count as 1 if it was generated in the language of the instruction (Greek, in this case), it belonged to the target category and it was not a repetition of a previous item uttered in the same sample. We categorised errors into three groups (following Wauters & Marquardt, 2018, with a few modifications): language choice errors, repetitions and category errors. Superordinate categories in the semantic task, proper nouns as well as unintelligible words were also excluded. For more detailed information regarding response codes concerning errors and exclusions, please see the project's OSF page (https://osf.io/dghsw/). Based on these scoring criteria, we calculated the number of correct responses produced on each task for each participant. Group means were calculated for each task.

We also analysed strategic processes (average cluster size; number of switches). For the analyses of clustering and switching, we followed closely Kosmidis et al. (2004), who had followed the scoring guidelines by Robert et al. (1998). Following them, we also did not count errors and exclusions when calculating clusters and switches. We calculated average cluster size and number of switches for each task separately.

For the semantic task, three or more consecutive words belonging to the same semantic subcategory were considered a semantic cluster (e.g., ελέφαντας-τίγρη-λιοντάρι-λεοπάρδαλη 'elephant-tiger-lion-leopard'). Semantic switches (i.e., number of transitions between clusters, including single words) were calculated by subtracting the total number of related words (i.e., all words forming a semantic cluster, 4 in the example above) from the total number of words produced and adding that to the number of semantic clusters. For the formal task, three or more consecutive words beginning with the same two letters and having the same sound (e.g., χάρισμα 'gift'-χαρά 'happiness'-χάδι 'caress') or two consecutive words that differed only in a vowel sound (e.g., χέρι 'hand'-χάρη 'favour') or words that were homophones (e.g., σύκο 'fig'-σήκω 'get up') were considered a formal cluster. Formal switches (i.e., number of transitions between clusters, including single words) were calculated by subtracting the total number of related words (i.e., all words forming a formal cluster, in the examples above 3, 2 and 2, respectively) from the total number of words produced and adding that to the number of formal clusters. For both tasks, single words that were not part of a cluster counted as one switch each. If a participant did not produce any clusters, the cluster size was set to 0. If there was no cluster, each word that was produced counted as a new switch. For detailed scoring rules, see appendix in Kosmidis et al. (2004, 171).

2.2. Data analysis

2.2.1. Number of correct responses

The bilingual participants generated a total of 3491 items on both tasks (in the three categories in the semantic task: animals, fruit and vegetables, and objects, and in the three letters in the formal task: *Χ*[chi], *Σ* [sigma] and *Α* [alpha]). Of these items, 140 (4.01%) were coded as errors or exclusions and were not included in the total number of correct responses. Repeated items were the most common type of exclusion (n = 80). The total number of correct items produced by bilinguals was 3351 items on both tasks.

https://doi.org/10.1017/S1366728924000671 Published online by Cambridge University Press

The monolingual participants generated a total of 3898 items on both tasks. Of these items, 182 (4.67%) were coded as errors or exclusions and were not included in the total number of correct responses. Repeated items were the most common type of exclusion (n = 111). The total number of correct items produced by monolinguals was 3716 items on both tasks.

The error rates across groups were similar. The most common error types in both groups were repetitions, amounting to almost half of the errors. Interestingly, language choice errors were very rare in the bilingual group (only 7 occurrences in total). Many bilinguals mentioned in the debriefing that they would think of English words, but apparently, they were able to suppress uttering them. For more detailed information regarding the numbers of errors and exclusions per response code and group, please see the project's OSF page (https://osf.io/dghsw/).

The total number of correct items produced for each category of each task for each group is shown in Table 1. Monolinguals produced more correct items in both semantic and formal fluencies, but the difference is more pronounced in semantic fluency. The greatest number of items was produced in the 'objects' condition by the monolingual group (M = 25.69; SD = 6.16) and the smallest number of items was produced in the Σ (sigma) condition by the bilingual group (M = 12.5, SD = 4.45).

We used R (R Core Team, 2019) and the lme4 package (Bates et al., 2015) to perform a generalised mixed-effects linear analysis of the effects between group (monolingual/bilingual) and task (semantic/formal) on the number of correct responses, specifying a Poisson family. The predictors were contrast coded and were modelled with glmer. First, we fitted a full model with group and task as fixed effects (with an interaction term) as well as age, proficiency in Greek (both scaled using the scale function in R) and gender and with random intercepts for subject. Then, we performed a likelihood ratio test of the full model with an interaction term against a model without the interaction term and the comparison proved nonsignificant ($\chi^2(1)$ = 1.8865, p = .1696). Including an interaction did not improve model fit, so we used the model without the interaction term for all subsequent analyses. Models were manually stepped-down (using likelihood ratio tests) from maximal models containing all factors and possible interactions to the 'best' model that only contained significant predictors or predictors that participated in significant interactions (Barr et al., 2013). The full model parameters of the best-fitting model are provided in Table 2.

Table 1. Total number of correct items per fluency type and group

	Bilingual		Monolingual		
Measure	*M*	*SD*	*M*	*SD*	Pairwise comparison
Semantic					
Animals	19.75	5.92	23.78	6.2	*** (p = .00746)
Fruit and vegetables	19.31	4.36	23.16	4.8	*** (p = .05962)
Objects	24.59	5.38	25.69	6.16	n.s.
Total	63.75	12.43	72.56	14.44	
Formal					
X (chi)	13.44	3.64	13.56	3.64	n.s.
Σ (sigma)	12.5	4.35	14.28	4.18	n.s.
A (alpha)	13.9	4.76	14.56	3.95	n.s.
Total	39.84	10.87	42.38	10.09	

Table 2. Estimates, standard errors, z values and p values of the best-fitting glmm for number of correct responses

	Estimate	Std. error	z value	Pr (<|z|)
Intercept (monolingual, semantic)	3.96539	0.02328	169.895	<0.001***
Group (bilingual)	−0.10510	0.04607	−2.281	0.0225 *
Task (formal)	−0.50557	0.02465	−20.506	<0.001***

As displayed in Table 2, there is a main effect of group (β = − 0.10510, z = −2.281, p = .02; with a small effect size as indicated by Cohen's d at 0.3) in that monolinguals produced more correct responses than bilinguals across the board. There is also a main effect of task (β = −0.50557, z = −20.596, p < .001; with a large effect size as indicated by Cohen's d at 2.18), as predicted, in that both groups produced more correct words in the semantic task than in the formal task. Neither age nor proficiency in Greek or gender proved to be significant predictors (age: β = 0.03547, z = 1.359, p = .1743; proficiency in Greek: β = 0.01192, z = 0.501, p = .6156; gender: β = 0.06942, z = 1.463, p = .1434).

In a second planned analysis, we looked at group differences within each task: the differences were big in the semantic task (β = −0.13067, z = −2.630, p < .008) and not as big in the formal task (β = − 0.06278, z = −1.133, p = .257), but this was only a nonsignificant trend as the group per task interaction had proven nonsignificant.

In a third planned analysis, we investigated possible differences among the Greek–English bilinguals in number of correct responses using the predictor variables of use of L1 (Greek) and length of residence, which were scaled using the scale function in R. However, none of these predictors proved to be significant (ps > .05).

In two exploratory analyses prompted by one of the reviewers, we conducted separate analyses for semantic and formal fluencies with individual categories as factors (semantic: animals, fruit and vegetables, objects | formal: chi, sigma, alpha). In semantic fluency, bilinguals significantly differed from monolinguals in the 'animals' (β = −0.184692; z = −2.676, p = .00746) and 'fruit and vegetables' (β = −0.180462; z = −2.595, p = .00946) categories, where bilinguals produced less items than monolinguals, but not in the 'objects' category (β = −0.03496, z = −0.53, p = .5962). In formal fluency, bilinguals did not differ significantly from monolinguals in any of the categories: chi (β = 0.01446, z = 0.161, p = .872), sigma (β = 0.13842, z = 1.538, p = .124) and alpha (β = 0.05130, z = 0.581, p = .562).[3]

2.2.2. Clustering and switching

The average cluster size and switching are displayed for each task and each group in Table 3. The monolinguals and the bilinguals are similar in average semantic clustering, while they differ in semantic switching, with bilinguals showing a smaller number of switches than monolinguals. We observe a similar picture in formal clustering, where the two groups are similar in average formal clustering and differ in formal switching, with bilinguals showing a smaller number of switches than monolinguals. All participants produced semantic clusters, but not all participants produced formal clusters.

[3]For a discussion of some of these data in a paper that explores participants' views on attrition and their feelings about their own use of Greek and English, see Lazaridou-Chatzigoga and Karatsareas (2022).

https://doi.org/10.1017/S1366728924000671 Published online by Cambridge University Press

Table 3. Average cluster size and number of switches per fluency type and group

	Bilingual	Monolingual
Semantic		
Average cluster size	3.22 (SD 0.96; range 1.33–5.67)	3.54 (SD 1.02; range 1.67–5.33)
Number of switches	43.06 (SD 8.56; range 31–60)	49.53 (SD 9.18; range 33–65)
Formal		
Cluster size	0.79 (SD 0.65; range 0–2.33)	0.625 (SD 0.5; range 0–2)
Number of switches	34.09 (SD 9.89; range 13–59)	37.86 (SD 9.7; range 19–59)

Thus, 12 out of the 64 participants (6 bilinguals; 6 monolinguals) produced no cluster in formal fluency.

We used R (R Core Team, 2019) and the lme4 package (Bates et al., 2015) to perform a mixed-effects linear analysis of the effects between group (monolingual/bilingual) and task (semantic/formal) on clustering. The predictors were deviation coded and were modelled with lmer. First, we fitted a full model with group and task as fixed effects (with an interaction term) as well as age, proficiency in Greek (both scaled using the scale function in R) and gender and with random intercepts for subject. Then, we performed a likelihood ratio test of the full model with an interaction term against a model without the interaction term and the comparison proved nonsignificant ($\chi^2(1) = 3.3189$, $p = .068$). Including an interaction did not significantly improve model fit, so we used the model without the interaction term for all subsequent analyses. The full model parameters of the best-fitting model are provided in Table 4.

As displayed in Table 4, there is no significant main effect of group ($\beta = -0.07813$, $t = -0.015$, $p = .609$) in that both monolinguals and bilinguals produced similar average cluster sizes in both tasks. There is a main effect of task ($\beta = -2.67188$, $t = -19.476$, $p < .001$; with a large effect size as indicated by Cohen's d at 3.26), in that both groups produced a greater average cluster size in the semantic task than in the formal task. Neither age nor proficiency in Greek or gender proved to be significant predictors (age: $\beta = 0.056599$, $t = 0.627$, $p = .5329$; proficiency in Greek: $\beta = 0.007415$, $t = -0.314$, $p = .7546$; gender: $\beta = 0.192542$, $t = 1.170$, $p = .2467$).

Our predictions were confirmed with respect to the formal task as we were expecting a similar performance in average cluster size in both bilinguals and monolinguals. Our predictions were not confirmed with respect to the semantic task although, as we anticipated a smaller average cluster size for bilinguals, but they produced equal cluster size.

In a second planned analysis, we investigated possible differences among the Greek–English bilinguals in average cluster size using the predictor variables of use of L1 (Greek) and length of residence, which were scaled using the scale function in R. However, none of these predictors proved to be significant (ps > .05).

We used R (R Core Team, 2019) and the lme4 package (Bates et al., 2015) to perform a mixed-effects linear analysis of the effects between group (monolingual/bilingual) and task (semantic/formal) on switching. The predictors were deviation coded and were modelled with lmer. First, we fitted a full model with group and task as fixed effects (with an interaction term) as well as age and

Table 4. Estimates, standard errors, df, t values and p values of the lmm for clustering

	Estimate	Std. error	df	t value	Pr (<|z|)
Intercept (monolingual, semantic, female)	2.04427	0.07590	62.00001	26.933	<0.001***
Group (bilingual)	−0.07813	0.15180	62.00001	−0.515	0.609
Task (formal)	−2.67188	0.13719	63.000002	−19.476	<0.001***

Table 5. Estimates, standard errors, df, t values and p values of the lmm for switching

	Estimate	Std. error	df	t value	Pr (<|z|)
Intercept (monolingual, semantic, female)	41.141	1.008	62.000	40.805	<0.001***
Group (bilingual)	−5.125	2.016	62.000	−2.542	0.0135*
Task (formal)	−10.312	1.184	63.000	−8.709	<0.001***

proficiency in Greek (both scaled using the scale function in R) and gender and with random intercepts for subject. Then, we performed a likelihood ratio test of the full model with an interaction term against a model without the interaction term, and the comparison proved nonsignificant ($\chi^2(1) = 1.317$, $p = .2503$). Including an interaction did not significantly improve model fit, so we used the model without the interaction term for all subsequent analyses. The full model parameters of the best-fitting model are provided in Table 5.

As displayed in Table 5, there is a main effect of group ($\beta = -5.125$, $t = -2.542$, $p = .0135$; with a small effect size as indicated by Cohen's d at 0.48) in that monolinguals produced a greater number of switches than bilinguals in both tasks. There is a main effect of task ($\beta = -10.312$, $t = -8.709$, $p < .001$; with a large effect size as indicated by Cohen's d at 1.07), in that both groups produced a greater number of switches in the semantic task than in the formal task. Neither age nor proficiency in Greek or gender proved to be significant predictors (age: $\beta = 0.7254$, $t = 0.605$, $p = .5478$; proficiency in Greek: $\beta = 0.8839$, $t = 0.817$, $p = .4172$; gender: $\beta = 1.4410$, $t = 0.658$, $p = .5128$).

We had predicted that we would find no differences in switching between the two groups. Nevertheless, monolinguals outperformed bilinguals in number of switches in both tasks.

In a second planned analysis, we investigated possible differences among the Greek–English bilinguals in switching using the predictor variables of use of L1 (Greek) and length of residence, which were scaled using the scale function in R. However, none of these predictors proved to be significant (ps > .05).

3. Discussion

This study set out to determine group differences in verbal fluency performance between a group of Greek–English bilinguals with Greek monolinguals. The second objective was to compare performance on each component of the verbal fluency task, the semantic and the formal. A third research objective was to include not only quantitative measures (i.e., number of correct responses), but also strategic processes (i.e., clustering, switching) in order to

https://doi.org/10.1017/S1366728924000671 Published online by Cambridge University Press

compare the strategies that the two groups use to generate words in this task and to understand the profile of bilinguals better.

To summarise the main findings, monolinguals performed significantly better in the correct number of responses across both tasks. This seems to be driven by the performance in the semantic task, where the differences between the two groups were big, and not by the performance in the formal task, where the differences were smaller. In a language attrition context that studies the L1 of the bilinguals involved, this is the first time that significant differences of this nature are reported. This is in line with Opitz (2011), who had also shown differences in the same direction, albeit of a nonsignificant nature, possibly due to the small sample size. As the bilingualism studies focus on the L2 of the bilinguals, we are not in a position to compare our results directly to this literature, but it is worth highlighting that the bilinguals that took part in our study performed worse than the monolinguals in both tasks in their L1.

A second important finding was that the interaction of the group per task was not significant. A second planned analysis showed that there is a trend for bilinguals to perform closer to the monolinguals in the formal task than in the semantic task. This would be in line with some research (see Luo et al., 2010) that shows that the performance of bilinguals does not differ from that of monolinguals on formal fluency. A closer look at the group per task interaction though shows that in most studies the interaction is nonsignificant (Paap et al., 2017). Paap et al. (2017) discuss and challenge the following four common assumptions that "predict a specific pattern of Group × Task interaction, namely, that the bilingual disadvantage in category fluency should be reduced, eliminated or possibly even reversed in a letter-fluency task". These assumptions are as follows: (a) formal fluency requires more executive control than semantic fluency, (b) bilinguals have enhanced abilities in executive function, (c) semantic fluency requires better lexical access skills than formal fluency and (d) bilinguals are worse than monolinguals in lexical access. In our study, we did not include executive function tasks, which would independently and directly test executive function, but our results show that the bilinguals were able to show a similar performance to that of the monolinguals in the formal task, as the trend showed.

Related to the different categories within each type of fluency, bilinguals and monolinguals were found to differ significantly only in the 'animals' and 'fruit and vegetables' categories, whereas in the 'objects' category they produced a similar number of items. This could be related to the fact that, as a reviewer suggests, 'animals' and 'fruit and vegetables' are categories with inherent organisational structure, whereas 'objects' is a large category with no obvious organisational structure. Participants seemed to structure their search for the 'objects' category based on what was found in their environment, which in all cases was a quiet café/office room. For instance, participant 12 produced "table, chair, tile, glass, computer, bottle etc.' all the objects they could see in their surroundings. This was not an option for the 'animals' and 'fruit and vegetables' categories as these were not available where the sessions took place, and participants had to rely on their memory in order to retrieve words in those categories.

On the strategies, with respect to average cluster size, we were expecting that bilinguals would show smaller cluster size in the semantic task given that we did not match the two groups on vocabulary and clustering uses more of the linguistic components. We were expecting possibly equal cluster size in the formal task. The results showed that there was no group difference in clustering in either task. The lack of a difference in clustering in the semantic task could be interpreted as an indication that the groups had similar vocabulary size, even though this was not independently measured.

With respect to switching, we were expecting that bilinguals would perform similarly to monolinguals in number of switches in both semantic and formal fluencies because the evidence for an increased executive control in bilinguals is far from settled (for two contrasting views, see Paap et al., 2017; Patra et al., 2020 among others). According to Patra et al. (2020), bilinguals are expected to switch more compared to monolinguals because it is a strategy that taps into executive control, and bilinguals are supposed to have better executive control than monolinguals. These assumptions are although challenged by Paap et al. (2017), who argue that a relatively better performance by a group on formal fluency does not constitute evidence of superior executive function unless independent measures are also included. In our study, monolinguals outperformed bilinguals in number of switches in both tasks. We think that these results challenge the assumption that we should expect bilinguals to rely more on switching due to their enhanced executive function abilities. Another reason for bilinguals showing more switching than monolinguals would have been the assumption that given that bilinguals have less lexical access, they would have to rely more on this strategy to generate words in the verbal fluency task. Nevertheless, this is also not supported by our data. Thus, our results show that monolinguals rely more on switching than bilinguals, who do not seem as able to tap into this strategy.

In conclusion, our results provide important new data that compare L1Greek-L2English late bilinguals to Greek monolinguals. We showed that monolinguals outperformed bilinguals in number of correct responses across both tasks with the semantic task driving the difference between the two groups. In clustering, both groups behaved similarly, while in switching monolinguals performed better than bilinguals. Our data attempted to address the gap in the literature of Greek but also in attrition studies, which usually do not go beyond the semantic task. By investigating both the semantic and formal tasks, we were able to get a fuller picture of the bilinguals' fluency. Although the group differences within each task were not significant, the bilinguals' performance was numerically closer to the monolinguals' performance on the formal task than on the semantic task. Looking at the results at the individual categories within each fluency type, we gained insight in the bilinguals' performance, which can inform future studies. Including at least one of the categories with inherent structure (e.g., 'animals' and 'fruit and vegetables') and one category without a clear structure (e.g., 'objects') can provide a better picture of the bilinguals' fluency. Finally, building on our results, future investigations could also include verbal fluency results in the bilinguals' L2 as well as their L1 in order to be able to compare the data to research stemming both from language attrition and bilingualism.

Data availability statement. The data that support the findings of this study are openly available at https://osf.io/dghsw/.

Acknowledgements. The authors would like to thank the editor as well as the two anonymous reviewers for their constructive comments. The authors are grateful to all participants for taking part in their studies. The authors would also like to thank Chara Triantafyllidou for data scoring. The project benefited from discussions with colleagues at the Research Unit (Experimental) Syntax and Languages (RUESHeL) at the Humboldt-Universität zu Berlin, where the first author was affiliated while conducting this study. The authors are also grateful to the audience at Queen Mary LingLunch. The project was supported by the DFG Gottfried Wilhelm Leibniz Preis 2014 to Artemis Alexiadou (AL 554/8-1).

Competing interest. The authors declare none.

https://doi.org/10.1017/S1366728924000671 Published online by Cambridge University Press

References

Aita, S., **Beach, J.**, **Taylor, S. E.**, **Borgogna, N. C.**, **Harrell, M. N.**, & **Hill, B. D.** (2019). Executive, language, or both? An examination of the construct validity of verbal fluency measures. *Applied Neuropsychology. Adult*, **26**(5), 441–451.

Ammerlaan, T. (1996). *'You get a bit wobbly…' – Exploring bilingual lexical retrieval processes in the context of first language attrition* [Ph.D. thesis]. *Katholieke Universiteit Nijmegen.*

Amunts, J., **Camilleri, J. A.**, **Eickhoff, S. B.**, **Patil, K. R.**, **Heim, S.**, **von Polier, G. G.**, et al. (2021). Comprehensive verbal fluency features predict executive function performance. *Scientific Reports*, **11**, 6929.

Barr, D. J., **Levy, R.**, **Scheepers, C.**, & **Tily, H. J.** (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**(3), 255–278.

Bates, D., **Maechler, M.**, **Bolker, B.**, & **Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Benton, A. L. (1968). Differential behavioral effects in frontal lobedisease. *Neuropsychologia*, **6**, 53–60.

Bialystok, E., **Craik, F. I. M.**, & **Luk, G.** (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, **21**, 522–538.

Birdsong, D., **Gertken, L. M.**, & **Amengual, M.** (2012). *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas. Available at https://sites.la.utexas.edu/bilingual/.

Cherciov, M. (2011). *Between attrition and acquisition: The dynamics between two languages in adult migrants* [Unpublished doctoral dissertation]. University of Toronto.

Filippetti, V. A., & **Allegri, R. F.** (2011). Verbal fluency in Spanish-speaking children: Analysis model according to task type, clustering, and switching strategies and performance over time. *The Clinical Neuropsychologist*, **25**, 413–436.

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**, 238–250.

Gollan, T. H., **Montoya, R. I.**, & **Werner, G. A.** (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, **16**, 562–576.

Gruenewald, P. J., & **Lockhead, G. R.** (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 225–240.

Keijzer, M. (2007). *Last in first out? An investigation of the regression hypothesis in Dutch emigrants in anglophone Canada* [Ph.D. thesis]. Vrije Universiteit Amsterdam.

Kormi-Nouri, R., **Moradi, A. R.**, **Moradi, S.**, **Akbari-Zardkhaneh, S.**, & **Zahedian, H.** (2012). The effect of bilingualism on letter and category fluency tasks in primary school children: Advantage or disadvantage? *Bilingualism: Language and Cognition*, **15**, 351–364.

Kosmidis, M. H., **Vlahou, C. H.**, **Panagiotaki, P.**, & **Kiosseoglou, G.** (2004). The verbal fluency task in the Greek population: Normative data, and clustering and switching strategies. *Journal of the International Neuropsychological Society*, **10**, 164–172.

Lazaridou, D., & **Karatsareas, P.** (2022). Language attrition and lived experiences of attrition among Greek speakers in London. *Languages*, **7**(4), 307.

Ljungberg, J. K., **Hansson, P.**, **Andrés, P.**, **Josefsson, M.**, & **Nilsson, L.-G.** (2013). A longitudinal study of memory advantages in bilinguals. *Plos One*, **8**(9), e73029.

Luo, L., **Luk, G.**, & **Bialystok, E.** (2010). Effects of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, **114**, 29–41.

Opitz, C. (2011). *First language attrition and second language acquisition in a second-language environment* [Ph.D. dissertation], Trinity College Dublin.

Paap, K. R., **Myuz, H. A.**, **Anders, R. T.**, **Bockelman, M. F.**, **Mikulinsky, R.**, & **Sawi, O. M.** (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology*, **29**, 89–112.

Patra, A., **Bose, A.**, & **Marinis, T.** (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition*, **23**(1), 204–218.

Portocarrero, J. S., **Burright, R. G.**, & **Donovick, P. J.** (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, **22**, 415–422.

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/.

Robert, P. H., **Lafont, V.**, **Medecin, I.**, **Berthet, L.**, **Thauby, S.**, **Baudu, C.**, & **Darcourt, G.** (1998). Clustering and switching strategies in verbal fluency tasks: Comparison between schizophrenics and healthy adults. *Journal of the International Neuropsychological Society*, **4**, 539–546.

Robert, P. H., **Migneco, V.**, **Marmod, D.**, **Chaix-Medicin, I.**, **Thauby, S.**, **Benoit, M.**, **Beau, C. H.**, & **Darcourt, G.** (1997). Verbal fluency in schizophrenia: The role of semantic clustering in category instance generation. *European Psychiatry*, **12**, 124–129.

Roberts, P. M., & **Le Dorze, G.** (1998). Bilingual aphasia: Semantic organization, strategy use, and productivity in semantic verbal fluency. *Brain and Language*, **65**(2), 287–312.

Rosselli, M., **Ardila, A.**, **Araujo, K.**, **Weekes, V. A.**, **Caracciolo, V.**, **Padilla, M.**, & **Ostrosky-Solí, F.** (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology*, **7**, 17–24.

Sandoval, T. C., **Gollan, T. H.**, **Ferreira, V. S.**, & **Salmon, D. P.** (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, **13**, 231–252.

Schmid, M. (2007). The role of L1 use for L1 attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 135–153). John Benjamins.

Schmid, M. S. (2011). *Language attrition*. Cambridge University Press.

Schmid, M. S., & **Jarvis, S.** (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, **17**(4), 729–748.

Shishkin, E., & **Ecke, P.** (2018). Language dominance, verbal fluency, and language control in two groups of Russian–English bilinguals. *Languages*, **3** (27), 1–17.

Spreen, O., & **Benton, A. L.** (1977). *Neurosensory center comprehensive examination for Aphasia*. University of Victoria, Neuropsychology Laboratory.

Troyer, A. K., **Moscovitch, M.**, & **Winocur, G.** (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, **11**, 138–146.

Vega-Mendoza, M., **West, H.**, **Sorace, A.**, & **Bak, T. H.** (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition*, **137**, 40–46.

Waas, M. (1996). *Language attrition downunder*. Peter Lang.

Wauters, L., & **Marquardt, T. P.** (2018). Category, letter, and emotional verbal fluency in Spanish-English bilingual speakers: a preliminary report. *Archives of Clinical Neuropsychology*, **28**, 1–14.

Yağmur, K. (1997). *First language attrition among Turkish speakers in Sydney*. Tilburg University Press.