Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 13 (3). 2023. 541-570
https://doi.org/10.14746/ssllt.28818
http://pressto.amu.edu.pl/index.php/ssllt

The use of lexical retrieval strategies by creative second language learners: A computational analysis of clustering and switching

Almudena Fernández-Fontecha ✉
University of La Rioja, Logroño, Spain
https://orcid.org/0000-0002-6557-0296
almudena.fernandez@unirioja.es

James Ryan
Carleton College, Northfield, Minnesota, USA
https://orcid.org/0000-0003-2238-2743
jryan@carleton.edu

Abstract

This study seeks to delve into the potential role of divergent thinking, a component of creativity, in second language learning. Specifically, we compare the use of lexical organization and production strategies of two groups of more and less creative EFL learners in year 12 through an automatic vectorial semantic analysis of their retrieval in three second language semantic fluency tasks. Consistent with previous research in the field of creativity, our findings indicate that the creative group retrieved more second language words than the less creative group. These words were less related to each other and to the stimulus categories than the words generated by the less creative group. While the creative participants' retrieval was based on an extensive use of switching, a slight but non-significant trend was found in the production of longer clusters by the less creative participants. These results yield interesting insights into the potential role of creativity in second language learning.

Keywords: creativity; lexical production; second language; clustering; switching

1. Introduction

As a multifaceted construct, creativity has been explored from cognitive, psychological, and social perspectives. In this study, we adopt a cognitive perspective by means of which creativity is understood as the individual's capacity to produce novel ideas and solutions to a problem (Guilford, 1959). Within the classical trait theory represented by the psychometric approach, which posits that the individual's cognitive traits can be measured via tests, two components of creativity are distinguished, namely *convergent thinking* and *divergent thinking*. Convergent thinking entails generating the most suitable answer to a problem from a variety of options (Cropley, 2006). By contrast, divergent thinking is viewed as the capacity of producing multiple possible answers. Guilford (1959) identified the following dimensions of divergent thinking*: fluency* (the capacity to generate many ideas or solutions to a given problem), *originality* (the ability to generate unique or unusual ideas or solutions), *elaboration* (the ability to develop and build upon ideas or solutions in a detailed and extensive way), and *flexibility* (the ability to change perspectives or approaches to produce a diverse set of ideas or solutions). These dimensions are traditionally tested through verbal and figural creativity performance, two different cognitive abilities of the creative individual (e.g., Kasirer et al., 2020). As a learnable skill, creativity is not fixed and can be developed through training and practice (Ritter & Mostert, 2017), which, among other benefits, might mitigate any potential temporal decline affecting creative thinking (Said-Metwaly et al., 2021).

Creativity has been extensively approached from a neurocognitive perspective mainly through network analyses of semantic memory, a division of long-term memory where world knowledge is stored (Schendan, 2012). In these studies, individuals with varying levels of creativity have been reported to exhibit differences in their semantic memory, which could account for higher associative fluency as well as more original responses based on more distant associations (Benedek et al., 2020; Bernard et al., 2019; Benedek & Neubauer, 2013; Kenett & Faust, 2019). In addition, creative people demonstrate a more flexible semantic memory network structure (Kenett et al., 2014; Kenett et al., 2018). However, with some exceptions (e.g., Ottó, 1998), the role of creativity has been widely overlooked in second language acquisition until very recently, when a new line of studies is emerging in this field (e.g., Albert & Kormos, 2011; Fernández-Fontecha, 2021; Fernández-Fontecha & Kenett, 2022; Mackey, 2020; Suzuki et al., 2022).

In the study of semantic memory in first languages, lexical organization, access, and retrieval have been generally approached through verbal fluency tasks. In these tasks, individuals have to recall in a given time as many words as they can that begin with a given letter or that fall under a specified semantic category. The standard metric addressed in research is the total count of words generated. However, these tasks have also facilitated the exploration of cognitive mechanisms that determine how semantically or phonemically related words are grouped together (clustering) within a subcategory, and how transitions between clusters occur when searching for a new subcategory (switching). Research on these strategies is abundant, for example, on speech pathologies (e.g., Bose et al., 2017; Park et al., 2022). In creativity research, clustering and switching mechanisms are consistent, respectively, with the operationalization of flexibility as the total number of different categories retrieved in idea generation (Guilford, 1967), and as the number of transitions between categories (e.g., Acar et al., 2019). Recent neurocognitive evidence on creative ability suggests that, while they are two separate cognitive processes that contribute to flexibility – with switching requiring a greater cognitive investment than clustering, since it involves exploring more remote categories (Mastria et al., 2021), they complement each other in idea generation (Ovando Tellez et al., 2022).

Troyer et al.'s (1997) manual approach has been traditionally applied in the analysis of clustering and switching processes. However, while still influential, this approach is subject to inter-rater variability, which cannot easily cope with the subjectivity imposed by semantic relationships beyond form or meaning. Recently, researchers have developed automatic methods to clustering identification, such as distributional or vectorial semantic methods, which have proved useful in addressing the issue of semantic distance (e.g., Ryan, 2013).

Based on a conceptualization of creativity as divergent thinking, and grounded within the psychometric tradition in creativity research, this paper attempts to determine whether the features of lexical organization and production typically identified in the first language (L1) of highly creative individuals (i.e., higher associative fluency, higher uncommonness, switching) are also found in the second language (L2). Specifically, it aims to identify and compare the second language word retrieval strategies (clustering and switching) used by more and less creative learners of English as a foreign language (EFL). For that purpose, automatic clustering analyses based on a distributional semantic model were conducted using the participants' responses to three semantic fluency tasks (red, games and entertainment, intelligence). As creativity can be developed through training and practice, evidence about these processes might inform the design of effective training programs for enhancing creative thinking in second language teaching.

2. Literature review

2.1. The study of second language lexical organization and retrieval

Research has implemented fluency tasks in the study of bilingual lexical organization and retrieval (e.g., Friesen et al., 2015; Luo et al., 2010). However, L2 lexical organization has been traditionally tackled through one-answer priming or association tests. Its study through verbal fluency tasks is relatively new in this field. These tasks have been used in a recent line of second language research to address, primarily, the number and type of words generated by second language learners according to their proficiency level or age (e.g., Jiménez Catalán, 2014). In addition, the study has investigated the impact of the stimulus category on word retrieval and revealed that concrete categories elicit higher lexical production compared to less concrete categories (Jiménez Catalán & Dewaele, 2017), probably because recalling concrete terms is easier (Fliessbach et al., 2006; Hell & de Groot, 1998). As regards word classes, nouns seem to be the most available grammatical category (e.g., Carcedo González, 2000).

Concerning clustering and switching, only a few studies have addressed these processes in second languages. By using a series of verbal fluency tasks, Tomé Cornejo (2015) found that a group of Spanish L1 speakers exceeded another group of Spanish L2 learners both in the number and size of clusters and number of switches. The highest proficiency learners performed more closely to native speakers on these measures. Palapanidi (2019) reached the same conclusion in a study that made use of a verbal fluency task and involved two groups of Greek learners with different proficiency levels of Spanish as a foreign language. In this study, the group of C1 learners was found to produce more words, clusters and switches than the group of B1 learners. More recently, Jiménez-Catalán (2022) has also examined the effect of age on EFL lexical output and the use of word association processes (clustering and switching) of children and adolescents with a fluency task (animals). According to the findings, teenagers retrieved more words and clusters. There were no statistically significant differences in the number of switches.

So far, in the identification of these organization and retrieval strategies, this type of studies has been mainly based on manual methods of clustering analysis, such as Troyer et al.'s (1997). While these manual techniques are successful in the identification of semantic relations, such as "X is a Y" (e.g., a lion is an animal) or "X is part of Y" (e.g., a finger is part of an arm), and word-family knowledge relations, such as those of derivatives or inflected forms, they fail to account for relationships between words beyond meaning or form, such as *volcano* and *heat,* or *car* and *road*. In this process, the concept of semantic distance is fundamental. *Semantic distance* refers to the relatedness beyond meaning or form of two words as visualized by their proximity to one another in a semantic space. Computational methods have recently been applied to clustering and switching analysis.

2.2. Automatic analysis of lexical organization

Two main trends in computational analysis have recently been applied to lexical organization in first languages and, less frequently, in second languages: (i) *network analysis* and (ii) *distributional* or *vectorial semantic analysis*. Network analysis models are based on graphs, that is, mathematical structures made up of nodes or vertices, which represent the entities in the network, and edges or links, which represent the relationships or connections between those entities. Widely used in the study of L1 semantic memory in psychology and cognitive neuroscience (e.g., Borge-Holthoefer & Arenas, 2010), in network analysis, the meaning of a word is deduced from the number and nature of the connections it shares with various other words. Findings of network analysis in second language have generally revealed a less dense and poorer organization of the L2 lexicon compared to the L1 lexicon (e.g., Borodkin et al., 2016). Beyond representing semantic memory, this method allows the study of behavioral and psychological consequences of variation in semantic memory. Thus, this line of inquiry has also found a more flexible network structure in the semantic memory of creative people that allows for the retrieval of uncommon (distant) concepts.

The second main method of computational analysis in this field is distributional or vectorial semantic analysis, which we use in this research. In light of serious issues with the traditional manual approach of clustering analysis (e.g., Troyer et al., 1997), in recent years a growing body of work is specifically exploring the application of methods from *distributional semantics* to the challenge of automatic semantic clustering analysis. In distributional semantics, semantic relatedness can be automatically computed by simply comparing word distributions in a corpus through a vector-space model of semantics. In other words, the relatedness of any two words can be visualized as the distance between those words in a semantic space. Owing to the ubiquity of this approach, in the world of computational linguistics, distributional semantics is sometimes called *vectorial semantics.*

The study of vectorial semantics advanced significantly with the advent of latent semantic analysis (LSA) near the end of the 1980s (Furnas et al., 1988; Landauer & Dumais, 1997). Importantly, LSA overcame certain issues of the classical vector-space model. While earlier projects applied LSA to the general analysis of verbal-fluency data (e.g., Elvevåg et al., 2010; Marklund et al., 2009), Ryan (2013) pioneered the use of LSA techniques in automatic clustering analysis. This author built a semantic space for the *animals* domain by subjecting approximately 5,000 Wikipedia articles about animals to LSA. More recently, based on artificial neural networks, the Word2Vec technique (Mikolov, Chen et al., 2013; Mikolov, Sutskever et al., 2013) has brought performance gains beyond LSA (e.g., Pereira et al., 2016).

The literature has already recognized the usefulness of vectorial models for analyzing L1 semantic memory. Examples include clinical studies examining the impact of head trauma on professional fighters (Ryan et al., 2013), dementia (Ryan, 2013), and schizophrenia (Roig et al., 2017). However, there remains a paucity of evidence of its implementation in languages different from the mother tongue. To our knowledge, only Fernández-Fontecha et al. (2021) have applied *vectorial analysis* to other languages. In this case, the same Word2Vec technique used in the present study allowed identification of different patterns of lexical production and organization strategies in L2 and L3 EFL learners. The implementation of this technique to the data of the current study is further explained in the method section.

2.3. Creativity in second language learning

A considerable amount of literature has reported a positive relationship between bi-/multilingualism and the speaker's creativity (Fürst & Grin, 2017; Kharkhurin, 2012). Although the nature of this relationship is still not fully clear, two possible reasons are commonly adduced to explain this advantage of bilinguals in creative thought: (i) the potential effect of enhanced executive functioning (Bialystok et al., 2012), which might also improve their performance in creative tasks (Kharkhurin, 2017), or (ii) the simultaneous access to multiple cultures and their conceptual representations, which may also contribute to their creativity (Kharkhurin, 2010; Leung et al., 2008). The majority of these studies have employed psychometric methods to investigate creativity in migrant populations.

In line with Van Dijk et al.'s (2019) suggestion of considering the role of environmental factors in exploring the relationship between creativity and language learning, we should note that formal academic foreign language learning, as it is the case of the present study, is a different scenario. Here, the speakers' linguistic profiles and their linguistic experiences tend to differ from those in the previous type of research. The available research into creativity in this context is still in its infancy, and, while in the studies above, creativity has been approached as a dependent variable of bi-/multilingualism, here it has mainly explored creativity as an independent variable. In other words, the focus has been on the impact that creativity might have on second language performance. Along this line of research, contradictory results have been reported about learners' creativity scores and L2 proficiency. While Ottó (1998) and Smith (2013) identified a positive correlation between these two variables, Albert (2006) did not find any relationship. Additionally, research has also demonstrated that creativity is somehow related to L2 fluency (e.g., Albert & Kormos, 2011; Fernández-Fontecha, 2021; Fernández-Fontecha & Kenett, 2022; McDonough et al., 2015; Suzuki et al., 2022). For instance, Albert and Kormos (2011) conducted a study on Hungarian secondary school students in which they found a positive relationship between creative fluency and the quantity of words recalled in two oral narrative tasks. However, they also observed a negative correlation between the number of words and originality. A positive relationship between divergent thinking and the number of words was also obtained by Suzuki et al. (2022), who made use of a picture narrative and an argumentative task to measure speech production in a group of Japanese-speaking learners of English at tertiary education. In Albert and Kormos (2011), the creativity scores were obtained through tasks based on the *Remote Associates Test* (Mednick, 1962) and the *Torrance Tests of Creative Thinking* (TTCT) (Torrance, 1990), among others. Suzuki et al. (2022) employed the *Alternate Uses Test* (Guilford, 1967) to assess divergent thinking, and a Japanese version of the *Remote Associates Test* (Mednick, 1962) to assess convergent thinking.

Finally, we should mention two recently published studies that have made use of a nearly identical sample of participants as the one used in this study. First, Fernández-Fontecha (2021) reported a positive relationship between EFL semantic fluency, measured via various fluency tasks different from the ones used here, and all the creativity constructs (i.e., fluency, flexibility, and originality). Second, Fernández-Fontecha and Kenett (2022) examined EFL learners' performance in two semantic fluency tasks in Spanish as L1 (animales) and English as L2 (animals) in terms of their creativity scores by using network science methods and a forward flow task, which quantified semantic distance based on latent semantic analysis (Beaty et al., 2021; Gray et al., 2019). Creative individuals were found to be more fluent in both languages, retrieved more distant responses, and displayed a more flexible and better organized semantic memory both in Spanish and English. The current research extends upon these studies by further exploring the relationship of creativity and second language fluency.

3. The study

3.1. Rationale

As explained above, higher associative fluency and the production of original ideas based on remote associations can be attributed to variations in the semantic memory of creative people. Both in effective and flexible generation of creative ideas and in any task involving semantic memory search, clustering and switching strategies are particularly relevant. The study of creativity as an individual variable in second language learning has received only cursory scholarly attention so far. By applying a vectorial semantic method in this study, we seek to identify the strategies used by two groups of EFL learners with different levels of creativity in various L2 semantic fluency tasks.

3.2. Research questions

We pursue the following research questions:

RQ1: Are there any differences in the EFL semantic fluency of the more creative (MC) and less creative (LC) groups of EFL learners regarding total types, unique and shared types, and mean number of words out of each participant's response?

RQ2: Are there any differences in the strategies of lexical organization and production used in EFL semantic fluency by the MC and LC groups?

3.3. Variables

The following variables will be addressed:

- *L2 semantic fluency* will be identified by means of three verbal fluency tasks, as we explain below. In particular, we will approach this variable in terms of *types* (different words or responses produced by each group of participants), and the *mean number of words* out of each participant's response.
- *Global creativity* will be measured through the PIC-J test (Artola et al., 2008), described below.
- Finally, *vectorial analysis* will help identify the following strategies: mean cluster size, mean chain size, mean cluster switches, mean chain switches, mean global and local pairwise relatedness, and mean category relevance.

A subtle variation on the notion of a semantic cluster is used in this study: the semantic *chain*. Following Troyer et al. (1997), a *cluster* is a group of words appearing contiguously in a test response such that each and every word is semantically related to each and every other word in the cluster. In a *chain*, however, only words appearing directly adjacently in the test response need to be related to one another. This differentiation is based on the idea that individuals tend to explore semantic fields locally rather than seeking out words associated with prior responses beyond the most recent one. The term *chain* is adopted from the second author's earlier work (Ryan, 2013; Ryan et al., 2013), but the general idea dates to at least Hills et al. (2012) and it has also been utilized in other more recent studies (Fernández-Fontecha et al., 2021; Linz et al., 2017; Pakhomov & Hemmy, 2014; Prud'hommeaux et al., 2017).

Based on these definitions, we will identify the *cluster/chain size* and *cluster/chain switch.* The former is understood as the number of members in the cluster or chain. In studies based on Troyer et al. (1997), at least two words are required to make a cluster. In this study, we consider singletons as one-word clusters/chains; hence, they also count toward the mean of this variable. The latter refers to a transition between two clusters/chains in a test response. Transitions between clusters/chains and singletons, and between singletons and other singletons are counted for this measure, thus, the number of cluster/chain switches in a response is simply the number of clusters/chains minus one. Group means for this metric were calculated by taking the mean across the number of cluster/chain switches generated by each participant (Ryan, 2013). In conclusion, the distinct cognitive phenomena of clustering (and chaining) and switching are therefore measured, respectively, by the mean cluster/chain size (including singletons), and the number of switches (i.e., number of clusters/chains minus one).

Global relatedness corresponds to the extent to which participants examine the category's semantic domain throughout their word retrieval. This metric is calculated by computing the mean across all pairwise relatedness scores given all the words in a test response, whether or not those words appeared adjacently. That is, for all possible pairwise combinations of the words in a test response, semantic relatedness is computed, and the mean of each of these scores is used to determine the final metric (Pakhomov et al., 2012; Ryan, 2013).

Local pairwise relatedness is calculated by computing the mean across all pairwise relatedness scores given all adjacent word pairs in the retrieval. This metric is aimed to gauge how far a respondent travels through the semantic space of a given category (Pakhomov et al., 2012; Prud'hommeaux et al., 2017; Thompson et al., 2013).

Category relevance assesses how much participants base their semantic search on the prompt category. We consider all the words in a test response, calculate the respective semantic relatedness scores between those words and the category prompt, and then average across all these scores. To our knowledge, only one earlier project has applied this metric (Nicodemus et al., 2014). In global, local relatedness, and category relevance, the score becomes lower as exploration of the semantic space increases. The reason for this behavior is that in fluency tasks the most related words will tend to be retrieved at the beginning. Hence, the more words are produced, the more likely it is that the response will include words that are less related to all the words in the retrieval, to the previous word, and to the stimulus category. In creativity studies, the most related words are considered more typical or less creative responses.

Considering the findings of prior research revealing a connection between fluency in first language and creativity, our main assumption is that the overall creativity score will also positively affect word production in the second language both quantitatively and qualitatively. Among others, semantic memory search strategies, such as clustering and switching, might contribute to this result. In accordance with findings in studies like Benedek and Neubauer (2013) or Benedek et al. (2020) in the first language, and in particular, Fernández-Fontecha and Kenett (2022) in the second language, the MC group will obtain lower values in local and global relatedness as well as in relation to the category, as they will produce more remote (weakly related) words than the LC group. In these results, some effect of the category could be expected.

4. Method

4.1. Participants

To obtain a valid sample of participants, we collected data from two schools with comparable socioeconomic status. The sample consisted of 34 12th grade students (25 males, nine females, aged between 17 and 18), who spoke Spanish as their L1 and had begun learning English as a foreign language at the age of four. The participants were roughly the same as in Fernández-Fontecha (2021) and Fernández-Fontecha and Kenett (2022). Due to outliers removal, the samples were not identical. Based on the PIC-J test (Artola et al., 2008), explained below, we identified the participants' scores for global creativity and divided the sample via a median split into two cohorts according to these scores: more creative (MC) (*N* = 17) and less creative (LC) participants (*N* = 17), as shown in Table 1.

Table 1 Descriptive statistics for global creativity

	M	*SD*	*n*	SE_M	Min	Max	Skewness	Kurtosis	*Mdn*
					Global creativity				
MC	120.94	29.95	17	7.26	94.00	211.00	1.83	2.94	112.00
LC	65.76	17.34	17	4.21	19.00	88.00	-1.02	1.06	65.00

Note. MC (more creative) participants, LC (less creative) participants

4.2. Instruments

4.2.1. Linguistic background questionnaire

Homogeneity in the linguistic profile of our sample of participants was controlled via a linguistic background questionnaire about their experience with EFL, which included inquiries about their mother tongue, other language(s) they knew, the additional EFL hours beyond the school curriculum, summer classes, travels abroad, or enrolment in any language program.

4.2.2. Semantic fluency tasks

In these tasks, the participants were asked to recall the very first words that came to their mind within each of the following stimuli: red, games and entertainment (games, henceforth), and intelligence. While in neuropsychology studies, verbal fluency is typically explored in one minute, in this study we used two minutes per task to avoid a possible ceiling effect on word retrieval. In this type of research, the category *animals* has been traditionally used (e.g., Bose et al., 2017; Rosselli et al., 2009). In the present study, we wanted to explore the participants' behavior in less cohesive (or less taxonomic) categories, which also differed in various traits, such as word class, frequency level, or concreteness value (i.e., the degree to which the concepts encoded by words refer to observable entities). All categories used are nouns, and, in addition, *red* can be also an adjective. Regarding concreteness, we followed Brysbaert et al.'s (2014) ratings. In the category *games and entertainment*, a mean value was tallied out of the values of each content word (*games - entertainment*). *Games* and *red* are the most concrete categories. The categories ranged from K1 to K3 frequency levels (VP-Compleat: BNC-COCA 1-25k) (Cobb, 2020b). Although *intelligence* is the less frequent word, it was not unknown to our Spanish speaking sample as it resembles its Spanish equivalent *inteligencia*. Finally, regarding the categories' productivity (i.e., which categories are the most and least productive in terms of the mean number word types (different words) produced by the learner), according to the two-tailed paired samples *t*-tests ($p > .05$) performed, no significant differences regarding mean number of words retrieved for each category were found. Table 2 shows the main features of the four categories. Table 3 displays the descriptive statistics for semantic fluency in each category.

Table 2 Characteristics of stimulus categories

	Word class	Frequency level (VP-Compleat: BNC-COCA 1-25k) (Cobb, 2020b)	Concreteness value (Brysbaert et al., 2014)	
Red	Noun and adjective	K-1	4.24	
Games	Noun	K-1	4.50	$M = 3.70$
Entertainment	Noun	K-2	2.92	
Intelligence	Noun	K-3	2.24	

Table 3 Descriptive statistics for semantic fluency per stimulus category

	M	*SD*	*n*	SE_M	Min	Max	Skewness	Kurtosis	*Mdn*
Red	11.76	3.89	34	0.67	4.00	22.00	0.27	0.20	11.00
Games	12.09	4.29	34	0.74	4.00	24.00	0.53	0.70	11.00
Intelligence	11.18	5.58	34	0.96	3.00	27.00	1.09	0.93	10.00

4.2.3. PIC-J test

This study follows the traditional psychometric approach and makes use of the PIC-J test developed by Artola et al. (2008), validated for measuring divergent thinking in Spanish secondary school students. It is based on the well-known *Alternate Uses Test* by Guilford (1967) and the *Torrance Tests of Creative Thinking* (TTCT) (Torrance, 1990). The global creativity score is a combination of the scores for verbal creativity and figural creativity from four tasks in Spanish. The first task requires the participants to write a story about a drawing depicting a boy and a girl by a lake. The second task asks them to list all the possible uses of a rubber pipe. The third task challenges them to imagine of the different things that could happen if the floor suddenly became elastic. The fourth task involves completing four drawings and giving them a title. Verbal creativity is computed from the scores for various divergent thinking dimensions (i.e., fluency, flexibility, and originality) measured through the first three tasks. Figural creativity is the result of computing the scores on originality, elaboration, title and extra features of the drawing in the fourth task. Additional information on the tasks and the scoring system can be found in Artola et al. (2008) and Barraca et al. (2010).

4.2.4. Oxford Quick Placement Test

The *Oxford Quick Placement Test* (OQPT) (version 2) (UCLES, 2001) consists of 60 multiple-choice and gap-fill questions about the use of English and it was used to determine the EFL proficiency level of the participants. It is administered on paper.

We performed a two-tailed independent samples t-test to compare the EFL proficiency of the MC and LC groups. The results showed a significant difference between the MC group (M = 34.12, SD = 6.87) and the LC group (M = 29.59, SD = 5.76) in their EFL proficiency score, $t(32) = 2.08$, $p = .045$. Yet, as displayed in Table 4, according to a Pearson correlation analysis, no significant correlation was identified between L2 proficiency and global creativity, nor between L2 proficiency and semantic fluency (i.e., words recalled by each participant in each stimulus category) in most categories. Only the relationship between EFL proficiency and *games and entertainment* was found significant.

Table 4 EFL proficiency: Pearson correlation results

	r	*p*
EFL proficiency – Global creativity	.31	.077
EFL proficiency – *red*	.19	.282
EFL proficiency – *games and entertainment**	.53	.001
EFL proficiency – *intelligence*	.25	.156

Note. $N = 34$, $*p < .001$

4.3. Data collection

After obtaining permission from the schools' headmasters and teachers for data gathering, the first author collected the data of the study in two sessions during class time in each high school. We kept the same test administration order in both data collection processes. The students filled out the background questionnaire in five minutes and completed a set of verbal fluency tasks in two minutes each. They did the PIC-J test in 28 minutes (seven minutes per task), and another set of verbal fluency tasks. Finally, they completed the OQPT test in 30 minutes. The verbal fluency tasks contained more stimulus categories than the ones analyzed in this study. No English language variables were assessed by the divergent thinking test and the linguistic background questionnaire, hence they were both administered in Spanish. The OQPT and the semantic fluency tasks, on the other hand, were administered in English as both tests evaluated English language. We used Spanish to provide the directions for each test.

4.4. Analyses

The first author followed Jiménez Catalán and Dewaele (2017) to edit and encode the responses for the fluency tasks in Excel spreadsheets. For the automatic clustering analyses of the word retrieval to the semantic fluency tasks, the second author used VFClust 0.1.1, a Python package based on latent semantic analysis (Ryan, 2013; Ryan et al., 2013). We excluded some of the terms in each category from the computation as they were not part of the model's vocabulary. In order to standardize spellings, the American English option was adopted. The edited retrieval was used to calculate the lexical organization variables for both groups. We implemented a pre-built Word2Vec model using the continuous bag-of-words technique explained in Mikolov, Chen et al. (2013) and Mikolov, Sutskever et al. (2013). The Google News dataset (with about 100 billion words and 300-dimensional vectors that represent over three million words and phrases) was used to train this model. In this vocabulary space, the distance between two positions can be represented by cosine similarity. Thus, the relatedness of any two words can be visualized as the distance between those words in the semantic space. Word2Vec applies cosine similarity to determine the similarity between any given pair of words to which it assigns a score within the range from -1 to 1. The closer to 1, the more related the words are. We followed Pakhomov and Hemmy (2014) and set a threshold for each category to decide if two words are related enough to be in the same cluster. Following Ryan (2013), the second author set the similarity threshold for each category at the 90th percentile, what resulted in the following threshold levels for each category: *red* (0.1993), *games* (0.2385), and *intelligence* (0.2337). Then we applied our Word2Vec model to the data.

Following the scoring system and instructions of the PIC-J test in Artola et al. (2008), we derived the standard score for global creativity based on scores on both verbal creativity (encompassing fluency, flexibility, and originality), and figural creativity.

The responses in the OQPT were manually scored and electronically encoded in Excel. SPSS (version 26.0; IBM Corp. Released, 2019) was used for descriptive statistics, a series of independent samples t-tests that compared findings in various measures, and some partial correlation analyses. Text Lex Compare v.4.3 New index Calc. (Cobb, 2020a) provided the total, unique and shared types (i.e., different words) retrieved in the EFL fluency tasks. The qualitative analysis of the retrieval per category was done by examining the words processed in Excel, which were assigned a relatedness score through the vectorial semantic analysis.

5. Results

5.1. Differences in EFL semantic fluency

The first research question of this study examined the differences in various aspects of the participants' L2 semantic fluency in terms of their creativity scores. The results of the descriptive statistics show that the LC group has lower values than the MC group for the mean number of words in all categories (see Table 5). Statistically significant differences were demonstrated by various independent samples t-tests (*red*: $t(32) = 3.61$, $p = .001$, $d = 1.24$, 95% CI [6.44, 1.79]; *games*: $t(32) = 3.34$, $p = .002$, $d = 1.14$, 95% CI [6.91, 1.67]; *intelligence*: $t(32) = 2.76$, $p = .009$, $d = 0.95$, 95% CI [8.38, 1.27]). A large effect size was found in all cases.

Since, as shown in Table 4, a significant relationship was identified between EFL proficiency and semantic fluency in one category (*games*), we wanted to further explore the relation between creativity scores and word retrieval through several partial correlation analyses where any potential impact of second language proficiency was controlled. The mean number of words produced in all categories was positively correlated to the learners' global creativity score (*red*: *r*(32) = 0.60, *p* = .001; *games: r*(32) = 0.44*, p* = .011; *intelligence*: *r*(32) = 0.62, *p* = .001).

Table 5 Descriptive statistics for semantic fluency per creativity group: mean number of words

	M	*SD*	*n*	SE_M	Min	Max	Skewness	Kurtosis	*Mdn*
				Red					
MC	13.82	2.51	17	0.61	10.00	18.00	-0.16	-1.15	14.00
LC	9.71	3.98	17	0.97	4.00	22.00	1.52	3.67	10.00
				Games					
MC	14.24	4.62	17	1.12	4.00	24.00	-0.01	0.44	14.00
LC	9.94	2.61	17	0.63	4.00	15.00	-0.48	0.36	10.00
				Intelligence					
MC	13.59	6.18	17	1.50	6.00	27.00	0.89	-0.43	11.00
LC	8.76	3.70	17	0.90	3.00	15.00	0.09	-1.24	8.00

With regard to total types, the MC group produced more (*M* = 119.67, *SD* = 12.01) than the LC group (*M* = 75, *SD* = 13.86) in all cases. The MC group was also more productive regarding unique types (*M* = 82, *SD* = 8.89) than the LC group (*M* = 37.33, *SD* = 9.81). In terms of the distribution of the unique vs. shared production within each group, the performance of the MC group was more regular than that of the LC group. Thus, while in the MC group the unique and shared types in all the categories stood in around a 68% - 31% relationship, in the LC group unique types were slightly superior to shared types in *intelligence* and *games,* and fewer unique than shared types were prompted for *red*. These differences in the proportion of unique and shared types in the LC group were smaller than in the MC group. Table 6 displays the absolute and relative (%) values of unique and shared types (i.e., types produced in both creativity groups) per category and creativity group identified by Text Lex Compare (Cobb, 2020a).

Table 6 Descriptive statistics for types by creativity group

	Total types	Unique types	Shared types	
		Red		
MC	108	75 (69.44%)	33	(30.55%)
LC	59	26 (44.06%)		(55.93%)
		Games		
MC	119	79 (66.38%)	40	(33.61%)
LC	83	43 (51.80%)		(48.20%)
		Intelligence		
MC	132	92 (69.70%)	40	(30.30%)
LC	83	43 (51.80%)		(48.20%)

The qualitative analysis suggested both some differences and similar patterns between categories and creativity groups. For instance, the category *red* triggered mainly co-hyponymic relations in the field of colors. Other shared words referred to objects or substances where the red color was typically found, for example, clothes, make-up, food, blood, and words indicating metaphorical relations with *red*, such as *love, danger*, or *temperature*. The longest chains in this category tended to be located at the beginning of both groups' retrievals. Yet, chain switches appeared to be more randomly distributed in the MC group, which generated more singletons from the beginning. The MC group traversed the semantic space of the category through singletons or chains that either served to increase the instances within the shared topics, for instance, food (e.g., *tomato – strawberry – cherry; wine – berry – apple)* or danger (e.g., *alert – signal; dangerous – stop; importance – danger*), or served to extend to additional topics, such as pain (e.g., *pain – damage; wrath – anger – stress*) or behavior (e.g., bad – *wrong – stop; incorrect – mistake*).

In relation to *games*, the shared retrieval mostly covered the fields of sports, computers and videogames, television, cinema, and other leisure activities, including the word *friends* referring to the main participants in these fields. The longest chains in both groups contained sports. The unique types in each group largely coincided on the shared topics. The MC participants produced more words within these topics, and they also incorporated superordinate words such as *free time* (four occurrences), *enjoy* (five occurrences) or *play* (three occurrences). In addition to nouns, as the most frequent word class, both groups generated verbs, but, while these referred to the same shared topics in the case of the LC group (e.g., *draw, clap, meet)*, the MC group used them to describe more specific actions within general sports or activities, for example, *bait, release, pull, push*. Furthermore, the MC group slightly excelled the LC group in the use of adjectives (e.g., *competitive, happy, cool)*. The switching distribution in this category exhibited a more random behavior compared to the previous category.

In the category *intelligence,* the shared production mainly covered the field of education, described through items referring to the action of *study*, the location or level of study (*university, school*), main elements indicating success or failure in achievement (*mark*), subjects (e.g., *maths, physics, chemistry*), positive adjectives (e.g., *clever, smart, good, excellent, hardworking*), abilities (*think, read*), the organ *brain*, or professions (*professor, scientist, doctor, engineer*). Interestingly, in the shared retrieval, the participants only mentioned subjects and professions related to formal sciences (*maths, computer*) and physical sciences (*physics*). Only some references to social sciences were made in each group; whereas in the MC group, four participants retrieved *philosophy*, in the LC group one participant produced the following chain: *poetry – science – psychology – history.* We should consider that these learners were studying subjects of the science and technology branch in secondary education. Following the tendency in previous categories, the chains were longer in the LC group, but the differences between the groups were smaller than in the previous cases. As in the previous categories, in *intelligence*, the MC group expanded the items within the shared semantic fields by adding others, such as *class* (four occurrences) or *grade* (three occurrences) via singletons or chains (e.g., *class – education*; *teacher – grade*). It also incorporated words related to other fields such as the arts (*art, draw, write*), inventions and creativity (*genius – inventor; creative – talent*), professions (*work – job; scientist – banker; doctor – medicine*), words related to the brain (*brain – learning; mind – brain; computers – brain; IQ – brain; neuron - problem solving - rational*), economy (*rich – millionaire; money*), or the domestic domain (*car – house; electronic; glasses*). Both groups retrieved approximately the same number of types referred to adjectives, however, the LC group tended to arrange them into two or three word chains (e.g., *silly – stupid – clever*; *stupid – dumb – clever*; *difficult – easier*; *smart – intelligent*; *easy – successful*; *clever – silly*; *good – excellent*; *clever – best*) and nine singletons. The MC group, on the other hand, produced four chains (*blue – white; fake – silly – think; fast – easy; clever – smart*) and 17 singletons. Finally, no evidence of a clear distributional pattern of switches was found in any of the groups.

5.2. Lexical organization and production strategies

The second research question explored the use of EFL lexical organization and retrieval strategies along with local and global semantic relatedness variables in each category. Tables 7, 8 and 9 present descriptive statistics for these measures. The results of the descriptive statistics show that the LC group has lower values than the MC group for the number of cluster switches in all categories. Based on the results of a set of independent samples t-tests, the observed differences were statistically significant at an alpha level of .05 (*red*: $t(32) = 3.61$, $p = .001$, $d = 1.24$, 95% CI [5.34, 1.49]; *games*: $t(32) = 2.94$, $p = .006$, $d = 1.01$, 95% CI [5.18, 0,94]; *intelligence*: $t(32) = 3.07$, $p = .004$, $d = 1.05$, 95% CI [6.75, 1.36]). Similarly, the LC group showed lower values than the MC group for the number of chain switches. These differences were also significant in all categories (*red*: $t(32) = 3.6$, $p = .001$, $d = 1.23$, 95% CI [5.34, 1.48]; *games:* $t(32) = 2.64$, $p = .013$, $d = 0.91$, 95% CI [4.9, 0.63]; *intelligence*: $t(32) = 3.30$, $p = .002$, $d = 1.13$, 95% CI [6.67, 1.57]). No significant differences were found in any of the categories regarding cluster and chain size. Concerning global and local pairwise relatedness, and category relevance, the findings pointed to lower values in the MC group in *red* and *games*. However, *red* was the only category for which these differences were statistically significant on global pairwise relatedness: $t(32) = -2.10$, $p = .043$, $d = -0.72$, 95% CI [-.002, -.14], and category relevance, $t(32) = -2.16$, $p = .038$, $d = -0.74$, 95% CI [-.01, -.18].

Table 7 Descriptive statistics: lexical organization and production strategies for *red* by creativity level

Red	*M*	*SD*	*n*	SE_M	Min	Max	Skewness	Kurtosis	*Mdn*
				Mean cluster size					
MC	1.68	0.46	17	0.11	1.13	2.80	1.13	0.30	1.56
LC	2.11	1.09	17	0.26	1.00	5.00	1.32	1.23	2.00
				Mean chain size					
MC	1.76	0.52	17	0.13	1.13	3.00	1.20	0.61	1.57
LC	2.47	2.10	17	0.51	1.00	10.00	2.90	8.07	2.00
				Mean number of cluster switches*					
MC	7.65	3.10	17	0.75	4.00	14.00	0.59	-0.62	7.00
LC	4.24	2.36	17	0.57	1.00	9.00	0.50	-0.79	4.00
				Mean number of chain switches*					
MC	7.35	3.18	17	0.77	3.00	14.00	0.51	-0.73	7.00
LC	3.94	2.28	17	0.55	0.00	8.00	0.34	-0.75	4.00
				Mean global pairwise relatedness*					
MC	0.18	0.08	17	0.02	0.12	0.34	0.97	-0.41	0.14
LC	0.25	0.12	17	0.03	0.08	0.48	0.24	-1.00	0.27
				Mean local pairwise relatedness					
MC	0.25	0.09	17	0.02	0.13	0.42	0.81	-0.69	0.22
LC	0.30	0.13	17	0.03	0.07	0.51	-0.19	-1.14	0.31
				Mean category relevance*					
MC	0.28	0.11	17	0.03	0.13	0.47	0.32	-1.25	0.25
LC	0.38	0.14	17	0.03	0.15	0.56	-0.43	-1.22	0.42

Note. **p* < .05

Table 8 Descriptive statistics: lexical organization and production strategies for *games* by creativity level

Games	*M*	*SD*	*n*	SE_M	Min	Max	Skewness	Kurtosis	*Mdn*
				Mean cluster size					
MC	1.53	0.25	17	0.06	1.11	2.00	0.35	-0.70	1.50
LC	1.69	0.74	17	0.18	1.10	4.00	2.02	3.71	1.40
				Mean chain size					
MC	1.69	0.34	17	0.08	1.11	2.50	0.60	0.03	1.67
LC	1.82	0.81	17	0.20	1.10	4.00	1.49	1.38	1.50
				Mean number of cluster switches*					
MC	8.12	3.55	17	0.86	2.00	16.00	0.25	-0.25	8.00
LC	5.06	2.41	17	0.58	0.00	9.00	-0.04	-0.43	5.00
				Mean number of chain switches*					
MC	7.47	3.56	17	0.86	1.00	15.00	0.08	-0.35	7.00
LC	4.71	2.44	17	0.59	0.00	9.00	0.25	-0.42	4.00
				Mean global pairwise relatedness					
MC	0.18	0.04	17	0.01	0.12	0.24	0.05	-1.11	0.18
LC	0.21	0.09	17	0.02	0.12	0.52	2.65	7.18	0.19
				Mean local pairwise relatedness					
MC	0.23	0.06	17	0.01	0.13	0.33	-0.30	-0.48	0.23
LC	0.26	0.08	17	0.02	0.15	0.49	1.25	1.72	0.25
				Mean category relevance					
MC	0.17	0.05	17	0.01	0.10	0.28	0.69	-0.60	0.15
LC	0.19	0.06	17	0.01	0.11	0.33	1.29	1.07	0.17

Note. **p* < .05

Table 9 Descriptive statistics: lexical organization and production strategies for *intelligence* by creativity level

Intelligence	*M*	*SD*	*n*	SE_M	Min	Max	Skewness	Kurtosis	*Mdn*
				Mean cluster size					
MC	1.30	0.15	17	0.04	1.09	1.50	0.05	-1.46	1.29
LC	1.32	0.27	17	0.07	1.00	1.86	0.59	-0.74	1.29
				Mean chain size					
MC	1.34	0.21	17	0.05	1.11	1.80	0.79	-0.33	1.29
LC	1.40	0.39	17	0.09	1.00	2.50	1.30	1.68	1.33
				Mean number of cluster switches*					
MC	9.41	4.85	17	1.18	3.00	20.00	0.87	-0.32	8.00
LC	5.35	2.50	17	0.61	2.00	9.00	0.06	-1.37	5.00
				Mean number of chain switches*					
MC	9.12	4.61	17	1.12	3.00	19.00	0.79	-0.41	8.00
LC	5.00	2.29	17	0.56	2.00	9.00	0.35	-0.86	5.00
				Mean global pairwise relatedness					
MC	0.14	0.03	17	0.01	0.08	0.20	0.06	-0.56	0.14
LC	0.14	0.04	17	0.01	0.05	0.23	-0.11	1.03	0.14
				Mean local pairwise relatedness					
MC	0.18	0.04	17	0.01	0.12	0.24	-0.10	-1.24	0.18
LC	0.18	0.07	17	0.02	0.04	0.30	0.02	-0.62	0.19
				Mean category relevance					
MC	0.08	0.02	17	0.00	0.05	0.12	0.65	-0.01	0.08
LC	0.07	0.02	17	0.01	0.01	0.10	-0.75	0.33	0.07

Note. *$p < .05$

6. Discussion

Fluency is one of the main components of creativity. The most creative individuals tend to produce a greater number of more uncommon responses in the L1. Differences in the L1 lexical organization of the creative mind have been reported in the literature that could explain this result. On the other hand, preliminary research has identified a positive connection between L2 performance and creativity, yet more research is still needed about this relationship. The present study compared the lexical organization and retrieval in a second language of two groups of EFL learners differing in their creativity levels. The word retrieval strategies adopted by each group were identified by means of a vectorial semantic analysis of the words retrieved in three verbal fluency tasks (*red*, *games*, and *intelligence*). These stimulus categories were homogeneous in productivity, that is, in the number of different words elicited, although they differed in frequency level and concreteness value. Creativity and L2 proficiency showed no significant correlation. Only fluency in one category was associated with L2 proficiency.

When it comes to RQ1, the results showed that the MC group elicited more total and unique word types than the LC group in all categories. The participants also produced more unique types than shared types, a result that occurred in a roughly similar distribution across categories. Fernández-Fontecha (2021) also reported this advantage in both total and unique types recalled in other semantic fluency tasks by a sample of participants roughly similar to the sample in the present study. In the LC group, the difference between the proportion of unique and shared types was smaller, and slightly more irregular across categories. Overall, these results might suggest that the most creative individuals are able to overcome the difficulties imposed by certain less predictable semantic categories (e.g., *intelligence*). Regarding individual behavior, the MC learners produced more words than the LC learners in all three stimuli. These results also align with Fernández-Fontecha (2021), which found a similar correlation between the creativity score, and all the dimensions of creativity, and the number of words retrieved in the four categories used in the study. These findings can also be generalized to other contexts and educational levels, as proved in previous research in the field of second language acquisition (e.g., Albert & Kormos, 2011; Suzuki et al., 2022). Although the main purpose of these studies was different to the purpose of the current study – they assessed oral production through a variety of narrative tasks, both identified a positive relationship between some creativity dimensions and the quantity of words recalled in the L2 tasks.

Our objective in RQ2 was to deepen our understanding of the findings from RQ1 by analyzing the participants' use of lexical organization and production strategies. The MC production was mainly based on cluster/chain switching strategies in all categories. No significant differences were found in any of the categories regarding cluster/chain size although the tendency pointed to the production of longer clusters and chains in the case of LC participants. A greater number of clusters and switches, as well as longer clusters, have been identified by research as more typical strategies used by L1 speakers and more proficient L2 learners (Palapanidi, 2019; Tomé Cornejo, 2015). Tomé Cornejo (2015) also found that switching had a stronger effect on L2 fluency, maybe to compensate for possible linguistic deficiencies. Studies in first languages have also detected a greater relevance of this switching strategy in semantic fluency (Troyer et al., 1997; Unsworth et al., 2011). On the one hand, in producing a larger number of switches, our MC participants replicated the behavior of native speakers and more L2 proficient learners. However, they did not produce longer clusters/chains, as these individuals did. Although we do not have full evidence of the cause of the adoption of the switching strategy by the MC group (i.e., whether it depends on creative ability, L2 proficiency, or other variables such as vocabulary size), we cannot fully affirm either that it happened as a compensatory strategy for EFL proficiency issues since proficiency did not correlate with creativity, and it only correlated with fluency in one semantic category. Consistent with earlier evidence on creativity research that showed that the most creative people have a less rigid and less structured semantic network than the less creative ones (Benedek et al., 2017; Kenett et al., 2014; Li et al., 2021), we could argue that the use of switching in the MC group might be related to a greater cognitive flexibility in the activation of strategic search processes.

In accordance with a higher production of unique types than the LC group, in terms of global and local pairwise relatedness and category relevance, our findings point to lower values in these measures in the MC group for *red* and *games*, although only in *red* the difference is statistically significant regarding global relatedness and category relevance. In all the categories, the MC group traversed the semantic space by delving further into the shared topics and expanding them to others related to the category. At least in *red* and *games*, this behavior resulted in a less related retrieval both globally, locally, and concerning the category. Overall, these findings seem to be consistent with previous research on creativity and fluency in the first language, which identifies a higher number of responses due to more distant semantic connections (e.g., Benedek & Neubauer, 2013; Benedek et al., 2020; Bernard et al., 2019; Kenett & Faust, 2019). They also match the results observed in Fernández-Fontecha and Kenett's (2022) study on L1 and L2, which revealed that more creative learners recalled more remote responses in both languages in two semantic fluency tasks (*animals* and *animales*). Additionally, in this study the network analysis revealed a more flexible semantic memory network structure in the case of the creative participants, a result that could help explain their performance in the production of remote associations.

Some effects of the category could be observed in various respects. The category defines the type of available words. For example, *red* triggered co-hyponymic relationships (i.e., other colors), or words that suggest metaphorical associations (e.g., behavior, love, danger). *Games* also triggered hyponyms (types of sports), and words related to leisure activities (e.g., cinema, TV, computers, music). Probably because it is a less concrete category, and it is less well-defined in terms of semantic relationships (e.g., it does not typically trigger hyponymic or meronymy relationships), *intelligence* elicited more varied semantic fields in both groups. This resulted in a greater number of switches, shorter clusters/chains, and lower global, local and category relatedness values than in the other two categories. Moreover, the category also seemed to influence the distribution of the clusters. Thus, only in *red,* the longest clusters were produced at the beginning of both groups' retrievals. All in all, these findings may imply that individuals with higher creativity levels employ the switching strategy more frequently than those with lower creativity as a mechanism to explore the category more efficiently while overcoming its complexity, which leads to an overall increase in word retrieval in both their first and second languages.

7. Conclusions

Creativity has been a long-neglected variable in the field of second language learning. Neuroscientific research on creativity has offered consistent evidence on the relationship between divergent thinking and fluency of ideas. Creative individuals tend to exhibit greater associative fluency and generate more original ideas based on remote associations. Their flexibility in the process of idea generation is based on the adoption of strategies like clustering and switching. By implementing an automatic vectorial semantic model to the analysis of L2 retrieval in three semantic fluency tasks, our study compared the L2 lexical organization and production strategies of two groups of EFL learners with different creativity levels. In general, statistically significant differences were found in the lexical organization of the L2 lexical-semantic network of the two groups of participants as well as in their lexical production strategies. In keeping with the evidence about the behavior of the creative individuals in the L1, in the L2 the creative group proved to produce more uncommon responses and was more fluent than the less creative group. The higher retrieval of uncommon responses in the creative group runs in parallel with their lower values in global, local, and category relatedness measures in the stimulus categories *red* and *games and entertainment*, although only differences in global and category relatedness measures were significant in the former. The superiority in L2 fluency performance is characterized by the extensive use of switching. This search mechanism allows the creative learners to keep producing words by shifting between clusters/chains. This proactive behavior occurs in all the categories, even in *intelligence*, whose lower levels of concreteness and frequency might be hindering the word retrieval in both groups. No differences were found regarding cluster and chain size, although the results indicate that the group of less creative individuals tends to exhaust the clusters/chains of highly similar items before continuing the exploration of the semantic space, resulting in longer clusters/chains.

In the present study, the extensive use of switching by the group of more creative learners cannot be attributed to a compensatory strategy for a possible deficiency in L2 performance as this variable had no relationship creativity and it only showed a significant correlation with fluency in one category. Yet, considering previous research, it might be due to a more flexible semantic network. However, we are not claiming either that proficiency does not ultimately have some effect on the outcomes. Due to prior evidence of a less dense and poorer organization of the L2 lexical-semantic network (Borodkin et al., 2016), we think that proficiency could be affecting both groups' L2 performance compared to their hypothetically more optimal behavior in the L1. Further work should investigate if the behavior found in the L2 is kept or attenuated in the L1. Even if the latter was the case, in the light of the provided evidence, differences are still likely to be found between individuals with different levels of creativity.

This study has some limitations that should be acknowledged and addressed in future research. One of the main limitations is the small sample size, which reduces the generalizability and validity of the findings. Further studies should aim to include larger and more varied samples of participants with different linguistic repertoires and L2 learning backgrounds. Additionally, further investigation should examine the influence of other potential interacting variables that were not considered in the present study, such as vocabulary size, openness to experience or working memory. This investigation could also include the analysis of the learners' performance in other types of written and oral L2 productive tasks.

Despite these limitations, this research certainly demonstrates the potential of employing analytical techniques from other disciplines in the field of second languages. We believe that this interdisciplinary line of research can open new avenues for future studies in this area. Some useful insights into the role of creativity in second language learning have been offered through our vectorial semantic analysis, which merit pedagogical intervention. Our research has shown that creative learners use a variety of cognitive strategies in performing L2 fluency tasks. Thus, since creativity and divergent thinking can be learned, it is important to design and implement systematic training programs across the curriculum and in the L2 classroom to foster these skills. In L2 teaching, any genuinely communicative approach, such as task-based language teaching, would naturally allow for the implementation of creativity strategies to enhance each of the creative components while learning the second language. Task design may incorporate creativity training (Lin, 2011; Ritter & Mostert, 2017) and provide explicit instructions about strategies that maximize divergent thinking and creativity prior to a task. Thus, both oral and written skills could be practiced by means of brainstorming, generation of creative ideas to a hypothetical situation, problem-solving or problem generation tasks, among others. Here, resources like the creative training guidelines for the L2 curriculum developed by LEAP (2014) would provide valuable ideas for creative task design.

Acknowledgement

This work was supported by the FEDER/Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación [grant number PGC2018-095260-B-I00].

References

Acar, S., Runco, M. A., & Ogurlu, U. (2019). The moderating influence of idea sequence: A re-analysis of the relationship between category switch and latency. *Personality and Individual Differences, 142,* 214-217. https://doi.org/10.1016/j.paid.2018.06.013

Albert, Á. (2006). Learner creativity as a potentially important individual variable: Examining the relationships between learner creativity, language aptitude and level of proficiency. In M. Nikolov & J. Horváth (Eds.), *Empirical studies in English applied linguistics* (pp. 77-98). P´ecs: Lingua Franca Csopor.

Albert, Á., & Kormos, J. (2011). Creativity and narrative task performance: An exploratory study. *Language Learning, 54*(2), 277-310. https://doi.org/10.1111/j.1467-9922.2004.00256.x

Artola, T., Barraca, J., Martín, C., Mosteiro, P., Ancillo, I., & Poveda, B. (2008). *PIC-J. Prueba de Imaginación creativa para jóvenes*. TEA Ediciones.

Barraca, J., Poveda, B., Artola, T., Mosteiro, P., Sánchez, N., & Ancillo, I. (2010). Three versions of a new test for assessing creativity in Spanish population (PIC-N, PIC-J, PIC-A). ECHA Conference, Paris, 7-9 July. https://web.teaediciones.com/Ejemplos/PIC.pdf

Beaty, R. E., Zeitlen, D. C., Baker, B. S., & Kenett, Y. N. (2021). Forward flow and creative thought: Assessing associative cognition and its role in divergent thinking. *Thinking Skills and Creativity, 41*, 100859. https://doi.org/10.1016/j.tsc.2021.100859

Benedek, M., & Neubauer, A. C. (2013). Revisiting Mednick's model on creativity-related differences in associative hierarchies: Evidence for a common path to uncommon thought. *The Journal of Creative Behavior, 47*(4), 273-289. https://doi.org/10.1002/jocb.35

Benedek, M., Jurisch, J., Koschutnig, K., Fink, A., & Beaty, R. E. (2020). Elements of creative thought: Investigating the cognitive and neural correlates of association and bi-association processes. *NeuroImage, 210,* 116586. https://doi.org/10.1016/j.neuroimage.2020.116586

Benedek, M., Kenett, Y. N., Umdasch, K., Anaki, D., Faust, M., & Neubauer, A. C. (2017). How semantic memory structure and intelligence contribute to creative thought: A network science approach. *Thinking & Reasoning, 23*(2), 158-183. https://doi.org/10.1080/13546783.2016.1278034

Bernard, M., Kenett, Y., Ovando-Tellez, M., Benedek, M., & Volle, E. (2019). Building individual semantic networks and exploring their relationships with creativity. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 138-144). Cognitive Science Society.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences, 16*(4), 240-250. https://doi.org/10.1016/j.tics.2012.03.001

Borge-Holthoefer, J., & Arenas, A. (2010). Categorizing words through semantic memory navigation. *European Physical Journal B, 74*(2), 265-270. https://doi.org/10.1140/epjb/e2010-00058-9

Borodkin, K., Kenett, Y. N., Faust, M., & Mashal, N. (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. *Cognition*, *156*, 60-70. https://doi.org/10.1016/j.cognition.2016.07.014

Bose, A., Wood, R., & Kiran, S. (2017). Semantic fluency in aphasia: Clustering and switching in the course of 1 min. *Internacional Journal of Language & Communication Disorders, 52*(3), 334-345. https://doi.org/10.1111/1460-6984.12276

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*(3), 904-911. https://doi.org/10.3758/s13428-013-0403-5

Carcedo González, A. (2000). *Disponibilidad léxica en español como lengua extranjera: el caso Finlandés: Estudio del nivel preuniversitario y cotejo con tres fases de adquisición*. Universidad de Turku.

Cobb, T. (2020a). *Text Lex Compare v.4.3 New index Calc.* [computer program]. Retrieved 6 Nov 2020 at https://www.lextutor.ca/cgi-bin/tl_compare/

Cobb, T. (2020b). VP-Compleat v.2.5. (BNC-COCA 1-25k) [computer program]. Retrieved 6 Nov 2020 at https://www.lextutor.ca/vp/comp/

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal*, *18*(3), 391-404. https://doi.org/10.1207/s15326934crj1803_13

Elvevåg, B., Foltz, P. W., Rosenstein, M., & DeLisi, L. E. (2010). An automated method to analyze language use in patients with schizophrenia and their first-degree relatives. *Journal of Neurolinguistics*, *23*(3), 270-284.

Fernández-Fontecha, A. (2021). The role of learner creativity in L2 semantic fluency. An exploratory study. *System, 103,* 102658, https://doi.org/10.1016/j.system.2021.102658

Fernández-Fontecha, A., Jiménez Catalán, R. M., & Ryan, J. (2021). Lexical production and organisation in L2 EFL and L3 EFL learners: A distributional semantic analysis of verbal fluency. *International Journal of Multilingualism*, 1-18. https://doi.org/10.1080/14790718.2021.2016770

Fernández-Fontecha, A., & Kenett, Y. N. (2022). Examining the relations between semantic memory structure and creativity in second language. *Thinking Skills and Creativity, 45,* 101067. https://doi.org/10.1016/j.tsc.2022.101067

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *Neuroimage, 32*(3), 1413-1421. https://doi.org/10.1016/j.neuroimage.2006.06.007

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

Furnas, G. W., Deerwester, S., Durnais, S. T., Landauer, T. K., Harshman, R. A., Streeter, L. A., & Lochbaum, K. E. (1988). Information retrieval using a singular value decomposition model of latent semantic structure. *Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 465-480). ACM.

Fürst, G., & Grin, F. (2017). Multilingualism and creativity: A multivariate approach. *Journal of Multilingual and Multicultural Development, 39*(4), 341-355. https://doi.org/10.1080/01434632.2017.1389948

Guilford, J. P. (1959). Three faces of intellect. *American Psychologist, 14*(8), 469-479.

Guilford, J. P. (1967). *The nature of human intelligence.* McGraw Hill.

Gray, K., Anderson, S., Chen, E. E., Kelly, J. M., Christian, M. S., Patrick, J., Huang, L., Kenett, Y. N., & Lewis, K. (2019). Forward flow: A new measure to quantify free thought and predict creativity. *American Psychologist, 74*(5), 539-554. https://doi.org/10.1037/amp0000391

Hell, J., & de Groot, A. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*(3), 193-211. https://doi.org/10.1017/S1366728998000352

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431-440.

IBM Corp. Released (2019). *IBM SPSS Statistics for Windows, Version 26.0.* Armonk, NY: IBM Corp.

Jiménez Catalán, R. M. (Ed.). (2014). *Lexical availability in English and Spanish as a second language*. Springer.

Jiménez Catalán, R. M. (2022). Lexical profiles of children and adolescent EFL learners in the semantic domain of animals. In V. De Wilde & C. Goriot (Eds.), *Second language learning before adulthood: Individual differences in children and adolescents* (pp. 133-154). De Gruyter Mouton. https://doi.org/10.1515/9783110743043

Jiménez Catalán, R. M., & Dewaele, J.-M. (2017). Lexical availability of young Spanish EFL learners: Emotion words versus non-emotion words. *Language, Culture and Curriculum*, *30*(3), 283-299.

Kasirer, A., Adi-Japha, E., & Mashal, N. (2020). Verbal and figural creativity in children with autism spectrum disorder and typical development. *Frontiers in Psychology, 11*(559238), 1-15. https://doi.org/10.3389/fpsyg.2020.559238

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience, 8*, 1-16. https://doi.org/10.3389/fnhum.2014.00407

Kenett, Y. N., & Faust, M. (2019). A semantic network cartography of the creative mind. *Trends in Cognitive Sciences*, *23*(4), 271-274.

Kenett, Y. N., Levy, O., Kenett, D. Y., Stanley, H. E., Faust, M., & Havlin, S. (2018). Flexibility of thought in high creative individuals represented by percolation analysis. *Proceedings of the National Academy of Sciences, 115*(5), 867-872. https://doi.org/10.1073/pnas.1717362115

Kharkhurin, A. V. (2010). Sociocultural differences in the relationship between bilingualism and creativity potential. *Journal of Cross-Cultural Psychology, 41*, 776-783.

Kharkhurin, A. V. (2012). *Multilingualism and creativity*. Multilingual Matters.

Kharkhurin, A. V. (2017). Language mediated concept activation in bilingual memory facilitates cognitive flexibility. *Frontiers in Psychology, 8*(1067), 1-16. https://doi.org/10.3389/fpsyg.2017.01067

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211-240. https://doi.org/10.1037/0033-295X.104.2.211

LEAP (Language Education & Partnerships Ltd. UK). (2014). *Creative thinking in literacy & language skills: Training guidelines.* Project reference: 2014-1-UK01-KA204-000081. Co-funded by the Erasmus + Programme of the European Union. http://www.creativethinkingproject.eu/training_guide/CTILLS_O1_TrainingGuide_EN.pdf

Leung, A. K.-y., Maddux, W. W., Galinsky, A. D., & Chiu, C.-y. (2008). Multicultural experience enhances creativity: The when and how. *American Psychologist, 63*(3), 169-181

Li, Y., Kenett, Y. N., Hu, W., & Beaty, R. E. (2021). Flexible semantic network structure supports the production of creative metaphor. *Creativity Research Journal*. https://doi.org/10.1080/10400419.2021.1879508

Lin, Y.-S. (2011). Fostering creativity through education: A conceptual framework of creative pedagogy. *Creative Education, 2*(3), 149-155. https://doi.org/10.4236/ce.2011.23021

Linz, N., Tröger, J., Alexandersson, J., & König A. (2017). Using neural word embeddings in the analysis of clinical semantic verbal fluency task. *IWCS 2017: 12th International Conference on Computational Semantics: Short papers*. https://doi.org/10.1515/applirev-2018-0114

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

Mackey, A. (2020). *Interaction, feedback and task research in second language learning: Methods and design*. Cambridge University Press.

Marklund, P., Sikstrom, S., Baath, R., & Nilsson, L-G. (2009). Age effects on semantic coherence: Latent semantic analysis applied to letter fluency data. *Third International Conference on Advances in Semantic Processing (SEMAPRO'09)* (pp. 73-76). IEEE.

Mastria, S., Agnoli, S., Zanon, M., Acar, S., Runco, M. A., & Corazza, G. E. (2021). Clustering and switching in divergent thinking: Neurophysiological correlates underlying flexibility during idea generation. *Neuropsychologia*, *158*, 107890. https://doi.org/10.1016/j.neuropsychologia.2021.107890

McDonough, K., Crawford, W. J., & Mackey, A. (2015). Creativity and EFL students' language use during a group problem-solving task. *TESOL Quarterly, 49*(1), 188-199. https://doi.org/10.1002/tesq.211

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220-232.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality*. Advances in Neural Information Processing Systems, 21*, 3111-3119.

Nicodemus, K. K., Elvevåg B., Foltz. P. W., Rosenstein, M., Diaz-Asper, C., Weinberger, D. R. (2014). Category fluency, latent semantic analysis and schizophrenia: A candidate gene approach. *Cortex*, *55*, 182-91. https://doi.org/10.1016/j.cortex.2013.12.004

Ottó, I. (1998). The relationship between individual differences in learner creativity and language learning success. *TESOL Quarterly, 32*(4), 763-773. https://doi.org/10.2307/3588011

Ovando-Tellez, M., Benedek, M., Kenett, Y. N., Hills, T., Bouanane, S., Bernard, M., Belo, J., Bieth, T., & Volle, E. (2022). An investigation of the cognitive and neural correlates of semantic memory search related to creative ability. *Communications Biology, 5*(1), 1-16. https://doi.org/10.1038/s42003-022-03547-x

Pakhomov, S. V. S., & Hemmy, L. S. (2014). A computational linguistic measure of clustering behavior on semantic verbal fluency task predicts risk of future dementia in the Nun Study. *Cortex*, 55, 97-106.

Pakhomov, S. V. S., Hemmy, L. S., & Lim, K. O. (2012). Automated semantic indices related to cognitive function and rate of cognitive decline. *Neuropsychologia 50*(9), 2165-2175.

Palapanidi, K. (2019). Manifestaciones de "clusters" y "switches" en el léxico disponible de aprendices griegos de ELE en diferentes niveles lingüísticos. *marcoELE. Revista de Didáctica Español Lengua Extranjera*, *28*, 1-10.

Park, J., Yoo, Y. R., Lim, Y., & Sung, J. E. (2022). Phonological and semantic strategies in a letter fluency task for people with Alzheimer's disease. *Frontiers in Psychology, 13*, 7959.

Pereira, F., Gershman, S., Ritter, S., & Botvinick, M. (2016). A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data. *Cognitive Neuropsychology*, *33*(3-4), 175-190.

Prud'hommeaux, E., Van Santen, J., & Gliner, D. (2017). Vector space models for evaluating semantic fluency in autism. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 32-37). Association for Computational Linguistics.

Ritter, S. M., & Mostert, N. (2017). Enhancement of creative thinking skills using a cognitive-based creativity training. *Journal of Cognitive Enhancement, 1*, 243-253. https://doi.org/10.1007/s41465-016-0002-3

Roig, J., Ichinose, M., Hong, S. J., Ryan, J. O., & Park, S. (2017). 64: A computational analysis of verbal fluency in schizophrenia. *Schizophrenia Bulletin*, *43*(suppl_1), 39. https://doi.org/10.1093/schbul/sbx021.103

Rosselli, M., Tappen, R., Williams, C., Salvatierra, J., & Zoller, Y. (2009). Level of education and category fluency task among Spanish speaking elders: Number of words, clustering, and switching strategies. *Aging, Neuropsychology, and Cognition, 16*(6), 721-744. https://doi.org/10.1080/13825580902912739

Ryan, J. O. (2013). *A system for computerized analysis of verbal fluency tests*. [Master's thesis, University of Minnesota]. https://users.soe.ucsc.edu/~jor/publications/ryanMastersThesis.pdf

Ryan, J. O., Pakhomov, S., Marino, S., Bernick, C., & Banks, S. (2013). Computerized analysis of a verbal fluency test. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (Vol. 2, pp. 884-889). Association for Computational Linguistics.

Said-Metwaly, S., Fernández-Castilla, B., Kyndt, E., Van den Noortgate, W., & Barbot, B. (2021). Does the fourth-grade slump in creativity actually exist? A meta-analysis of the development of divergent thinking in school-age children and adolescents. *Educational Psychology Review*, *33*, 275-298. https://doi.org/10.1007/s10648-020-09547-9

Schendan, H. E. (2012). Semantic memory. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (2nd ed., pp. 350-358). Academic Press. https://doi.org/10.1016/B978-0-12-375000-6.00315-3

Suzuki, S., Yasuda, T., Hanzawa, K., & Kormos, J. (2022). How does creativity affect second language speech production? The moderating role of speaking task type. *TESOL Quarterly*. https://doi.org/10.1002/tesq.3104

Smith, C. A. (2013). Student creativity and language performance. In N. Sonda & A. Krause (Eds.), *JALT 2013 Conference Proceedings* (pp. 285-297). JALT.

Thompson, G., Kello, C., & Montez, P. (2013). Searching semantic memory as a scale-free network: Evidence from category recall and a Wikipedia model of semantics. *Proceedings of the Annual Meeting of the Cognitive Science Society 35*, 3533-3538. https://escholarship.org/uc/item/9mk6c2sc

Tomé Cornejo, C. (2015). *Léxico disponible: Procesamiento y aplicación a la enseñanza de ELE* [Master's thesis, Universidad de Salamanca]. https://goo.gl/vzwjjD

Torrance, P. (1990). *The Torrance tests of creative thinking*. Scholastic Testing Service.

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*(1), 138-146.

UCLES. (2001). *Quick placement test. Version 2. Photocopiable.* Oxford University Press.

Unsworth, N., Spillers, G. J., y Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology, 64*(3), 447-466. https://doi.org/10.1080/17470218.2010.505292

Van Dijk, M., Kroesbergen, E. H., Blom, E., & Leseman, P. P. (2019). Bilingualism and creativity: Towards a situated cognition approach. *The Journal of Creative Behavior, 53*(2), 178-188.