Published in final edited form as:
Clin Linguist Phon. 2009 June ; 23(6): 431–445. doi:10.1080/02699200902839800.

Semantic verbal fluency in two contrasting languages

SEIJA PEKKALA[1], **MIRA GORAL**[2,3,4], **JUNGMOON HYUN**[2,3], **LORAINE K. OBLER**[2,3,5], **TIMO ERKINJUNTTI**[6], and **MARTIN L. ALBERT**[3,5]

[1]Department of Speech Sciences, University of Helsinki, Helsinki, Finland [2]Program in Speech-Language-Hearing Sciences, CUNY, Graduate Center, New York, NY, USA [3]Medical Research Service, VA Boston Healthcare System, Boston, MA, USA [4]Lehman College, CUNY, New York, NY, USA [5]Department of Neurology, Boston University School of Medicine, Boston, MA, USA [6]Department of Neurology, University of Helsinki, Helsinki, Finland

Abstract

This cross-linguistic study investigated Semantic Verbal Fluency (SVF) performance in 30 American English-speaking and 30 Finnish-speaking healthy elderly adults with different cultural and linguistic backgrounds. Despite the different backgrounds of the participant groups, remarkable similarities were found between the groups in the overall SVF performance in two semantic categories (animals and clothes), in the proportions of words produced within the first half (30 seconds) of the SVF tasks, and in the variety of words produced for the categories. These similarities emerged despite the difference in the mean length of words produced in the two languages (with Finnish words being significantly longer than English words). The few differences found between the groups concerned the types and frequencies of the 10 most common words generated for the categories. It was concluded that culture and language differences do not contribute significantly to variability in SVF performance in healthy elderly people.

Keywords

Semantic verbal fluency; culture; language; cross-language; older adults

Introduction

There is a growing body of neuropsychological literature demonstrating that assessment tools generated in the context of a given culture may not be appropriate or accurate when assessing individuals from another culture (e.g. Nell, 2000; Rosselli and Ardila, 2003). The basis for these concerns is typically *not* the assumption that people who belong to varying ethnicities or cultures are fundamentally different, cognitively, but rather that the instruments used in neuropsychology may be culturally biased and, therefore, norms obtained from certain populations might not be useful when assessing other populations. This approach is substantiated by a variety of studies that have documented cultural differences in performance on a variety of cognitive tasks (see Park, Nisbett, and Hedden, 1999; Nisbett and Masuda, 2003 for reviews).

Correspondence: Dr Seija Pekkala, Department of Speech Sciences, University of Helsinki, PO Box 9 (Siltavuorenpenger 20 A), FIN—00014, Finland. Tel: + 358-0-191 29676; +358-50-3564 653. Fax: +358-9-191 29341. seija.pekkala@helsinki.fi.

***Declaration of interest*:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Cultural differences have been observed in tasks requiring attentional and working memory skills (e.g. WAIS-III/WAIS-R Digit Span tests). For example, Hedden, Park, Nisbett, Ji, Jing, and Jiao (2002) have shown that Chinese participants performed significantly better on the Digit Span test than American participants, whereas Boone, Victor, Wen, Razani, and Pontón (2007) have indicated that Caucasian participants scored similar to Asian participants, but significantly higher than African-American and Hispanic participants. Moreover, Dick, Teng, Kempler, Davis, and Taussig (2002) have found that Hispanic participants performed significantly worse on the Digit Span test than African-American, Caucasian, Chinese, and Vietnamese participants. Boone et al. (2007) have demonstrated that processing speed on the Trail Making Test (Part A) was significantly faster among Caucasians than African-Americans, and Hedden et al. (2002) have shown that when performing same-different comparisons of digit strings, Chinese participants were significantly faster than American participants.

Tasks assessing semantic memory have also been found to be susceptible to the effect of culture. For example, cultural differences can appear in the way semantic information is organized and used for tasks requiring reasoning and categorization: East Asians have been argued to show a tendency to organize objects in a relational way, whereas European Americans tend to group them in a categorical (taxonomic) way (Gutchess, Yoon, Luo, Feinberg, Hedden, Jing, Nisbett, and Park, 2006; Ji, Zhang, and Nisbett, 2004). For instance, when asked to group together two out of three pictured objects (e.g. ‘cow’, ‘chicken’, and ‘grass’), Chinese children grouped ‘cow’ and ‘grass’ together because cows eat grass, whereas American children grouped ‘cow’ and ‘chicken’ together because they were both animals (Chiu, 1972). The differences in semantic organization can also manifest as different recall and recognition strategies between these cultures (Park et al., 1999).

Cultural differences in the organization of semantic categories have been found not only between cultures that are distant from each other, such as East Asian and Western cultures (Lin, Schwanenflugel, and Wisenbaker, 1990; Ji et al., 2004; Yoon, Feinberg, Hu, Gutchess, Hedden, Chen, Jing, Cui, and Park, 2004; Gutchess et al., 2006), but also within Western cultures (Brown and Davies, 1976; Hasselhorn, Jaspers, and Hernando, 1990). The familiarity or centrality of an item in a culture is an important factor in determining the semantic category structure in children (Hasselhorn et al., 1990; Lin et al., 1990), younger adults (Brown and Davies, 1976; Schunn and Vera, 2004; Yoon et al., 2004), and older adults (Yoon et al., 2004). Nevertheless, the most typical exemplars of a semantic category appear consistent across cultures, at least for some categories, such as natural kinds (e.g. farm animals) (Hasselhorn et al., 1990; Yoon et al., 2004).

These reported cultural differences might result from a multitude of factors, such as differences in social, historical, economical, environmental, and educational systems. They might be due to the tested populations’ relative familiarity with specific items on the test and with the types of manipulation required by a given test rather than to superior performance of a particular group. Moreover, recent studies have indicated that early life experiences can explain cognitive performance in adulthood. For example, poor socioeconomic status can be associated with poorer cognitive functions at late adulthood (Kaplan, Turrell, Lynch, Everson, Helkala, and Salonen, 2001; Luo and Waite, 2005). Performance differences could also result from language-related differences (e.g. word length, number of exemplars of a given semantic category). Identifying measures that could transcend these differences and be used across cultures and languages would be beneficial to our ability to assess different populations with the same neuropsychological measures and to generalize across subject populations. In this study, we set out to assess whether the Semantic Verbal Fluency (SVF) task could be a useful assessment tool across cultures.

The SVF task, also known as the list generation task, in which participants are instructed to produce as many words as possible that belong to a certain semantic category (e.g. animals, fruit, vegetables) within a given time (typically 60 seconds), is a widely used method to investigate lexico-semantic and cognitive abilities in monolingual (e.g. Binetti, Magni, Cappa, Padovani, Bianchetti, and Trabucchi, 1995; Chan and Poon, 1999; Benito-Cuadrado, Esteba-Castillo, Böhm, Cejudo-Bolivar, and Peña-Casanova, 2002; Kavé, 2005; Goral, Spiro, Albert, Obler, and Connor, 2007) and bilingual individuals (Roberts and Le Dorze, 1997; Rosselli, Ardila, Salvatierra, Marques, Matos, and Weekes, 2002), and cross-linguistically (Kempler, Teng, Dick, Taussig, and Davis, 1998). The SVF task offers a fast and easily administered way to observe processes of search and access to information stored in semantic memory (Roberts and Le Dorze, 1997). This task has proved particularly useful in the study of language changes and cognitive changes associated with healthy ageing (e.g. Troyer, Moscovitch, and Winocur; 1997; Acevedo, Loewenstein, Barker, Harwood, Luis, Bravo, Hurwitz, Aguero, Greenfield, and Duara, 2000; Troyer, 2000; Brucki and Rocha, 2004; Kavé, 2005; Van der Elst, Van Boxtel, Van Breukelen, and Jolles, 2006) and in different neuropathological conditions, such as dementia (e.g. Rosser and Hodges, 1994; Suhr and Jones, 1998; Traykov, Baudic, Raoux, Latour, Rieu, Smagghe, and Rigaud, 2005; Pekkala, Albert, Spiro, and Erkinjuntti, 2008). Findings suggest that fewer correct responses and greater individual variability in the total number of words produced characterize the performance of older adults when compared to their younger counterparts (e.g. Troyer, 2000; Kosmidis, Vlahou, Panagiotaki, and Kiosseoglou, 2004; Kavé, 2005; Van der Elst et al., 2006). Our review of studies that used the SVF task with healthy older adults illustrates a substantial variability in the number of words produced in 1 minute among the studies, with means ranging from 8–22 words in the category of animals (see Table I), yet it is unclear whether these differences are culturally based.

Both similarities and differences have been found in studies that employed the SVF task across populations. For example, Acevedo et al. (2000) showed that a similar number of words were produced for the category of fruit by native English speakers and native Spanish-speaking immigrants living in Florida, but that differences emerged in the number of items produced for the categories of vegetables and animals. Namely, the English-speaking participants produced more words than the Spanish-speaking participants in the category of vegetables and the Spanish speakers produced more words than the English speakers in the category of animals. The authors associated these cross-linguistic differences in word production to differences in the distinct early living environments and exposure to different types of vegetables and animals between the English and Spanish speakers.

Environment has also been demonstrated to affect certain aspects (e.g. the frequency and type of vocabulary used) of performance on the SVF task. Kempler et al. (1998) showed that the most frequent animals named by African-American, White, Chinese, Hispanic, and Vietnamese speakers living in southern California were the same (i.e. ‘dog’, ‘cat’, ‘horse’, and ‘elephant’). However, they also found that the frequency of other animal names demonstrated regional and cultural differences among the groups: ‘ox’ and ‘buffalo’ were common for the Vietnamese speakers, ‘donkey’ for the Spanish speakers, ‘rat’ for the Chinese speakers, and ‘giraffe’ for the English speakers. Environment can also influence the vocabulary of speakers of the same language living in different areas of the same country (Brown and Davies, 1976; Brucki and Rocha, 2004). For example, Brucki and Rocha (2004) demonstrated that Brazilian-Portuguese speaking participants living in a city frequently produced ‘horse’ and ‘dog’, whereas those living in rural and rainforest areas produced more instances of ‘jaguar’ and ‘monkey’.

The role of education level in performance on the SVF task has also been inconclusive, with some studies demonstrating that those individuals with higher education produced significantly more responses for the task than those with fewer years of formal education (Kempler et al.,

1998; Chan and Poon, 1999; Acevedo et al., 2000; Benito-Cuadrado et al., 2002; Mathunarnath, George, Cherian, Alexander, Sarma, and Sarma, 2003; Kosmidis et al., 2004; Van der Elst et al., 2006). By contrast, other studies have shown only little impact of education on SVF performance (Harrison, Buxton, Husain, and Wise, 2000; Kavé, 2005; see also Troyer, 2000).

Language-specific factors, such as word length, could affect SVF performance because the access to and production of longer words requires more time and memory capacity than shorter words (Cheung and Kemper, 1994; Hedden et al., 2002). For example, Kempler et al. (1998) found that, even when the effects of age and education were controlled for, Spanish-speaking participants produced the fewest words and Vietnamese-speaking participants the most. The authors attributed these results to the differences in word length between the languages because animal names in Vietnamese tend to be one-syllable words whereas most animal names in Spanish are two-syllable words or longer. The word-length effect is consistent with the assumption that the total number of correct words produced for a category is related to the speed of word retrieval (Troyer et al., 1997; Vinogradov, Kirkland, Poole, Drexler, Ober, and Shenaut, 2003). Most of the words, moreover, are typically produced during the first half minute of the task, with fewer words being produced in the second half of the task (e.g. Rosen, 1980; Ober, Dronkers, Koss, Delis, and Friedland, 1986; Butters, Granholm, Salmon, Grant, and Wolfe, 1987; Brucki and Rocha, 2004).

In sum, variability in performance on the SVF task has been associated with a number of variables, including cultural differences. However, despite differences in performance among participant groups, there might be sufficient evidence for comparable results that could merit the usefulness of the SVF task as an assessment tool.

In the present study we consider speakers of two distinct cultures and languages by comparing SVF performance of two groups of healthy elderly people living in different environments. The participants were native speakers of two unrelated languages (American English and Finnish) that, in addition to a number of linguistic characteristics, differ from each other in word-length. Many more English words than Finnish words are monosyllabic, so, overall, a list of the same number of words in the two languages should have more syllables in the Finnish than in the English version (Elderton, 1949; Lehtonen, Sajavaara, and May, 1977; Hakulinen, Vilkuna, Korhonen, Koivisto, Heinonen, and Alho, 2004). We hypothesized that if cultural and language-related factors contribute to SVF performance, SVF performance would show differences between the participant groups. The different cultural backgrounds would result in differences in the most frequently produced words and the variety of vocabulary (i.e. a type/token ratio). The effect of language, in the form of a difference in word length, would be demonstrated not only as a difference in the total number of words but also as the number of correct words produced for the first half of the task (30 seconds). By contrast, if the factors that contribute to variation in performance on the SVF task predominately concern cognitive abilities and strategies that transcend cultural and linguistic differences we would find comparable performance in the two participant groups.

Method

Participants

The participants of this study were 30 monolingual English-speaking (16 female and 14 male) and 30 monolingual Finnish-speaking participants (14 female and 16 male) with the mean age of 67.5 years (SD=5.0) in the American English group and 66.7 years (SD=5.5) in the Finnish group (see Table II). The English-speaking participants comprised community-dwelling adults from the greater Boston area who participated in the Language in the Ageing Brain research project. The Finnish-speaking participants came from the capital city of Helsinki and its

surrounding cities participating in the Helsinki Ageing Study (Ylikoski, 2000). Participants had no history of neurological or psychiatric disease or substance abuse. Mean age and gender did not differ between the groups, but there was a statistically significant difference between the educational level of the American group (mean formal years of education=15.0, SD=2.0) and the Finnish group (mean formal years of education=9.7, SD=3.3) ($F(1, 58)=55.442$, $p<.001$) due to the relatively greater impact of World War II on the education opportunities of Finns and Americans of the older generation. The length of formal education required in Finland in the 1930s, 1940s, and 1950s was 6 years, relative to 10–12 years in Massachusetts. All participants gave their informed consent. The studies were approved by the IRBs of Boston University, the Boston VA Healthcare System, and the Ethics Committee of the Helsinki University Central Hospital.

Procedures and analyses

As part of a larger testing session, the participants were asked to produce as many responses as possible for two semantic categories: animals and clothes. Sixty seconds were allotted for each category. Participants' responses were audio-taped and transcribed verbatim. We first compared the two groups' overall SVF performance on the total number of correct words produced (excluding repetitions and words that did not belong to the target category). To reflect the diversity of the vocabulary used by the speakers of each language, we used the type/token ratio (number of unique words/total number of words) and counted the proportion of the 10 most frequently occurring words for animals and clothes. Finally, in order to look at the effect of language on SVF performance, we counted the number of syllables of each word for the mean length of words and the percentage of words produced within the first half (30 seconds) of the task. The variables were first calculated for each group in each semantic category and then combined for both categories.

All protocols were coded for correct responses and errors (perseverations and outside-category words) by native speakers of each language. Twenty per cent of each data set was analysed by independent raters, also native speakers of the respective languages. Inter-rater reliability was 99% for both the animals and the clothes in the American English data as well as the Finnish data.

Statistical analyses

Because the two groups differed significantly in their mean years of formal education (see above), we employed analysis of covariance (ANCOVA), with the years of education as the covariate. A chi-square (χ^2) test was administered to examine the differences between the groups in the type/token ratio as well as production of different types of exemplars for the categories. In all comparisons, an alpha level of .05 was used as the cut-off for the statistical significance. Mean (M), standard deviation (SD), and range were chosen for the measures of central tendency of the distribution and the variation in the data.

Results

When education was co-varied for,[1] ANCOVA revealed no significant differences between the groups in the number of correct words produced, the type/token ratio, and the number of words produced in the first 30 seconds (see Table III). This was true despite the finding that the two participant groups differed significantly with respect to the mean length of the words they produced. These findings were consistent in each semantic category and when the semantic categories were combined.

[1]When education was not accounted for, the two groups differed in the total number of correct words produced when the two categories were collapsed (ANOVA: $F(1, 58)=4.75$, $p<.05$) but not in any of the other measures.

The 10 most frequent exemplars produced for the animal category (Table IV) represented 34.7% of all the words produced by the English-speaking group and 38.2% of those produced by the Finnish-speaking group. There was no difference between these percentages ($\chi^2(1)=.736$, p=.391). The distribution of the most frequent animals indicated that seven out of the 10 most frequent animal words were similar for the two participant groups, and that they included pets, farm animals, and zoo animals. However, there were differences in the order of the most frequent members of the category between the groups: While most of the English speakers produced pets followed by zoo animals, most of the Finnish speakers produced exemplars of farm animals and pets, followed by zoo animals.

The proportion of the 10 most frequently produced words of the total number of words produced for the category of clothes was the same for both groups: 36.4% among the English speakers and 42.7% among the Finnish speakers (Table IV; ($\chi^2(1)=1.92$, p=.166). The most frequent words in each participant group consisted mainly of indoor and outdoor clothes. Except for 'scarf' in the English-speaking group and 'hame' ('skirt') in the Finnish-speaking group, the most frequent words were similar for the two participant groups.

Discussion

In this study, we examined SVF performance of two groups of elderly adults, monolingual speakers of American English and monolingual speakers of Finnish who differed in their language and cultural backgrounds. We employed a cross-language comparison between the two participant groups to examine SVF performance on two distinct semantic categories, animals and clothes. Despite the difference in socio-cultural backgrounds and the lexical characteristics of the languages (reflected here as the difference between the mean syllable length of words) of the two participant groups, we found that their SVF performance was very similar. This finding suggests that the factors that contribute most to variation in performance on the SVF task might be cognitive, rather than cultural or linguistic.

In contrast to previous findings of cross-language differences in the total correct word produced (Kempler et al., 1998; Acevedo et al., 2000), we found—after correcting for differences in formal education—no difference between the participant groups in the total correct words produced in either individual category or when the categories were combined. Even when education was not controlled for, the only significant difference we found between the participant groups was for the number of total words produced on the two categories combined ($p<.05$). Of course, because of the absence of truly education-matched participant groups, caution needs to be taken when interpreting these findings. Further study with larger participant groups might reveal group differences.

Our finding that education did not play a significant role in the SVF performance of the participant groups, apart from the total number of combined words, can be related to the observation that years of formal education do not directly reflect intellectual abilities (e.g. Ostrosky-Solís, Ardila, and Rosselli, 1999; Manly, Schupf, Tang, and Stern, 2005). For example, Bolla, Gray, Resnick, Galante, and Kawas (1998) demonstrated that verbal intelligence was more related to the ability to generate words for the SVF task than was the level of formal education achieved. Socio-economic and historical issues are likely to confound years of education in our cohort as many of the Finnish participants were prevented from attaining extended formal education (e.g. due to World War II). Thus their aptitude and the richness of their lexicon may be higher than the length of their formal education suggests (Chan and Poon, 1999; Kavé, 2005; Van der Elst et al., 2006). Furthermore, despite the relative lower range of education of the Finnish-speaking participants in our study, their total number of words found (mean=18.9) was within the range of previous studies (e.g. 8–22, see Table I). Moreover, it is worth noting that the upper-limit score of 22 found in our review was reported

in another study conducted with Finnish-speaking elderly participants who had formal education ranging from 6–12 years (Kontiola, Laaksonen, Sulkava, and Erkinjuntti, 1990), not by participants with the highest possible education range. Even though we did not set out to directly test the role of education in SVF performance, our findings add to the inconclusive effect of formal education and suggest that additional cross-linguistic research is needed to explore this matter. Employing measures of literacy, rather than years of formal education, as suggested by Manly et al. (2005), might prove fruitful, potentially resolving the inherent problem of differing typical length of formal education across cultures.

In addition, we found similar type/token ratios for the two participant groups for both individual and combined categories, suggesting that, overall, items belonging to the two semantic categories we tested were retrieved to the same extent by the English-speaking and Finnish-speaking participants. Furthermore, we found no differences between the groups in the proportion of the 10 most frequent words produced for the tasks. The only minor differences we found between the groups involved the types and the frequencies of the 10 most common words produced for animals and clothes. This difference could be explained by the different experience the two groups had with animals, owing in part to their socio-cultural backgrounds. In particular, the Finns, in general, grew up in more rural and northern areas, as well as a socio-economically less privileged society than their English-speaking American counterparts in our study, where they were frequently exposed to farm animals by experience. This may explain the prominent role of the production of farm animals in the SVF performance of the Finnish speakers. As far as the English speakers are concerned, their most frequently produced words contained more exemplars that can be found in zoos than those of their Finnish peers. This, in turn, can be an indication of the more urban backgrounds and early and more frequent exposure to zoo animals of the Boston-area participants. Our finding is in line with the study of Brucki and Rocha (2004) who indicated qualitative differences in the SVF responses between rural and urban populations. Recall, however, that we did not observe such differences for the category of clothes, for which we found virtually the same items produced by the two participant groups. This is not consistent with earlier studies in which the categories involving personal items, such as clothing (Hasselhorn et al., 1990) and footwear (Yoon et al., 2004), have been found to reflect sensitivity to the effects of culture. The similarities between our participant groups in the clothing category might be related to the cold weather shared by the environment of both groups. Taken together, these findings suggest that even though the overall SVF performance is similar between participants with different cultures and languages, specific portions of the lexicon can be affected by experience and cultural differences, and that cross-cultural differences may be category-specific.

Despite the fact that in our data the Finnish words were on average significantly longer than the English words produced, the Finnish-speaking participants produced a similar number of words for each category as the English-speaking participants. We interpret this to be consistent with the assumption that, at least at the group level, retrieval speed per se does not substantially contribute to performance on the SVF task, unlike the suggestions in some previous studies (e.g. Troyer et al., 1997; Vinogradov et al., 2003). Admittedly, the current study does not employ a direct measure of speed, but we plausibly hypothesize that if speed of production mattered, people speaking languages with longer words would produce fewer items either in the whole minute or within the first 30 seconds. That the Finnish-speaking group, while producing longer words than the English-speaking group, did not produce fewer words for the tasks contrasts with the findings of Kempler et al. (1998), in which Spanish speakers generated significantly longer and fewer words for the SVF task than did Vietnamese speakers. Furthermore, the word length difference did not yield a difference in the distribution of responses across the 60 seconds of the task: as in previous studies (Rosen 1980; Ober et al., 1986; Butters et al., 1987; Crowe, 1998; Brucki and Rocha, 2004), most of the words in the current results were generated during the first half of the tasks, in both language groups.

Our finding that no significant differences were observed between the two participant groups, despite their differences in culture and language, diverges from studies showing cross-cultural differences in cognitive functions (Chiu, 1972; Brown and Davies, 1976; Hasselhorn et al., 1990; Lin et al., 1990; Dick et al., 2002; Ji et al., 2004; Yoon et al., 2004; Gutchess et al., 2006; Boone et al., 2007). A potential explanation may be that culture-related differences are task- (or even category-) specific and that the semantic organization of the lexicon is mostly comparable across cultures. Alternatively, Park and colleagues (Park et al., 1999; Park and Gutchess, 2006) have hypothesized that cross-cultural differences in basic cognitive processes are smaller when older—as compared to younger—adults from various cultures are tested. It would be of interest to conduct cross-language comparisons between younger adults and older adults to examine age-related vs culture-related changes in the task. Furthermore, additional cross-linguistic research could address the interaction between education and SVF performance by comparing people varying in length of education among a number of countries sharing similar socio-cultural backgrounds. Also, as in the present study, the analysis of the SVF task could be extended beyond the total number of words to investigate the role of education on other cognitive aspects of the SVF performance.

Our findings suggest that the SVF task can be a powerful assessment tool across study populations. We do not suggest that cultural differences need can be discounted in neuropsychological studies, but rather that this specific task may serve as a relatively culture-free tool. The results of this study have clinical relevance demonstrating that, at least for studies involving older adults, the SVF task may reveal cognitive processes that are common to individuals from different cultural and language backgrounds.

Acknowledgements

This study was supported in part by The Finnish Cultural Foundation and The Academy of Finland (project number 213023, S. Pekkala). It was also supported in part by the Clinical Science Research and Development Service, US Department of Veterans Affairs, by Grant AG14345 from the National Institute on Aging. We thank Rossie Clark-Cotton, Avron Spiro, Kit Brady, Rebecca Williams, Elaine Dibbs, Jordan Awerbach, Becky Brown, and Jason Cohen for their contributions to this project. We would like to thank Kaisa Jauhiainen and Michael Krieger for analysing the data for the inter-rater reliability of this study. We thank Daniel Kempler for his insightful comments on an earlier version of this paper and we thank the reviewers of the manuscript. We also thank all our participants.

References

Acevedo A, Loewenstein DA, Barker WW, Harwood DG, Luis C, Bravo M, Hurwitz DA, Aguero H, Greenfield L, Duara R. Category fluency test: normative data for English- and Spanish-speaking elderly. Journal of the International Neuropsychological Society 2000;6:760–769. [PubMed: 11105466]

Benito-Cuadrado MM, Esteba-Castillo S, Böhm P, Cejudo-Bolívar J, Peña-Casanova J. Semantic verbal fluency of animals: a normative and predictive study in a Spanish population. Journal of Clinical and Experimental Neuropsychology 2002;24:1117–1122. [PubMed: 12650236]

Binetti G, Magni E, Cappa SF, Padovani A, Bianchetti A, Trabucchi M. Semantic memory in Alzheimer's disease: an analysis of category fluency. Journal of Clinical and Experimental Neuropsychology 1995;1:82–89. [PubMed: 7608305]

Bolla KI, Gray S, Resnick SM, Galante R, Kawas C. Category and letter fluency in highly educated older adults. The Clinical Neuropsychologist 1998;12:330–338.

Boone KB, Victor TR, Wen J, Razani J, Pontón M. The association between neuropsychological scores and ethnicity, language, and acculturation in a large patient population. Archives of Clinical Neuropsychology 2007;22:355–365. [PubMed: 17320344]

Brown WP, Davies GM. Studies in word listing: testing for group differences in category norms. Irish Journal of Psychology 1976;iii:87–120.

Brucki SMD, Rocha MSG. Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. Brazilian Journal of Medical and Biological Research 2004;37:1771–1777. [PubMed: 15558183]

Butters N, Granholm E, Salmon DP, Grant I, Wolfe J. Episodic and semantic memory: a comparison of amnesic and demented patients. Journal of Clinical and Experimental Neuropsychology 1987;9:479–497. [PubMed: 2959682]

Carew TG, Lamar M, Cloud BS, Grossman M, Libon DJ. Impairment in category fluency in ischemic vascular dementia. Neuropsychology 1997;11:400–412. [PubMed: 9223144]

Chan AS, Poon MW. Performance of 7- to 95-year-old individuals in a Chinese version of the category fluency test. Journal of the International Neuropsychological Society 1999;5:525–533. [PubMed: 10561933]

Cheung H, Kemper S. Recall and articulation of English and Chinese words under memory preload conditions. Language and Speech 1994;37:147–161.

Chiu L-H. A cross-cultural comparison of cognitive styles in Chinese and American children. International Journal of Psychology 1972;7:235–242.

Crowe SF. Decrease in performance on the verbal fluency test as a function in time: evaluation in a young healthy sample. Journal of Clinical and Experimental Neuropsychology 1998;20:391–401. [PubMed: 9845165]

Dick, MB.; Teng, EL.; Kempler, D.; Davis, DS.; Taussig, IM. The cross-cultural neuropsychological test battery (CCNB): effects of age, education, ethnicity, and cognitive status on performance. In: Ferraro, FR., editor. Minority and cross-cultural aspects of neuropsychological assessment. Swets & Zeitlinger; Lisse: 2002. p. 17-41.

Elderton WP. A few statistics on the length of English words. Journal of the Royal Statistical Society. Series A (General) 1949;112:436–445.

Fillenbaum GG, Heyman A, Huber MS, Ganguli M, Unverzagt FW. Performance of elderly African American and white community residents on the CERAD neuropsychology battery. Journal of Clinical and Experimental Neuropsychology 2001;7:502–509.

Goral M, Spiro A, Albert ML, Obler LK, Connor LT. Change in lexical-retrieval skills in adulthood. The Mental Lexicon 2007;2:215–238.

Gutchess AH, Yoon C, Luo T, Feinberg F, Hedden T, Jing Q, Nisbett RE, Park DC. Categorical organization of free recall across culture and age. Gerontology 2006;52:314–323. [PubMed: 16974103]

Hakulinen, A.; Vilkuna, M.; Korhonen, R.; Koivisto, V.; Heinonen, T.; Alho, I. ISO suomen kielioppi. Suomalaisen Kirjallisuuden Seura; Helsinki: 2004.

Harrison JE, Buxton P, Husain M, Wise R. Short test of semantic and phonological fluency: normal performance, validity and test-retest reliability. British Journal of Clinical Psychology 2000;39:181–191. [PubMed: 10895361]

Hart S, Smith CM, Swash M. Word fluency in patients with early dementia of Alzheimer type. British Journal of Clinical Psychology 1988;27:115–124. [PubMed: 3395734]

Hasselhorn M, Jaspers A, Hernando M-D. Typizitätsnormen zu zehn Kategorien für Kinder von der Vorschule bis zur vierten Grundschulklasse. Sprache & Kognition 1990;9:92–108.

Hedden T, Park DC, Nisbett R, Ji L-J, Jing Q, Jiao S. Cultural variation in verbal versus spatial neuropsychological function across the life span. Neuropsychology 2002;16:65–73. [PubMed: 11853358]

Ji L-J, Zhang Z, Nisbett RE. Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. Journal of Personality and Social Psychology 2004;87:57–65. [PubMed: 15250792]

Kaplan GA, Turrell G, Lynch JW, Everson SA, Helkala E-L, Salonen JT. Childhood socioeconomic position and cognitive function in adulthood. Interenational Journal of Epidemiology 2001;30:256–263.

Kavé G. Phonemic fluency, semantic fluency, and difference scores: normative data for adult Hebrew speakers. Journal of Clinical and Experimental Neuropsychology 2005;27:690–699. [PubMed: 16019645]

Kempler D, Teng EL, Dick M, Taussig IM, Davis DS. The effects of age, education, and ethnicity on verbal fluency. Journal of the International Neuropsychological Society 1998;4:531–538. [PubMed: 10050357]

Kontiola P, Laaksonen R, Sulkava R, Erkinjuntti T. Pattern of language impairment is different in Alzheimer's disease and multi-infarct dementia. Brain and Language 1990;38:364–383. [PubMed: 1693304]

Kosmidis MH, Vlahou CH, Panagiotaki P, Kiosseoglou G. The verbal fluency task in the Greek population: normative data, and clustering and switching strategies. Journal of the International Neuropsychological Society 2004;10:164–172. [PubMed: 15012836]

Lehtonen, J.; Sajavaara, K.; May, A. Spoken English: The Perception and Production of English on a Finnish-English Contrastive Basis. Gummerus; Jyväskylä: 1977.

Lin P-J, Schwanenflugel PJ, Wisenbaker JM. Category typicality, cultural familiarity, and the development of category knowledge. Developmental Psychology 1990;26:805–813.

Luo Y, Waite LJ. The impact of childhood and adult SES on physical, mental, and cognitive well-being in later life. Journal of Gerontology: Social Sciences 2005;60B:S93–S101.

Manly JJ, Schupf N, Tang M-X, Stern Y. Cognitive decline in literacy among ethnically diverse elders. Journal of Geriatric Psychiatry and Neurology 2005;18:213–217. [PubMed: 16306242]

Mathuranath PS, George A, Cherian PJ, Alexander A, Sarma SG, Sarma PS. Effects of age, education and gender on verbal fluency. Journal of Clinical and Experimental Neuropsychology 2003;25:1057–1064. [PubMed: 14566579]

Nell, V. Cross-Cultural Neuropsychological Assessment: Theory and Practice. Lawrence Erlbaum; Mahwah, NJ: 2000.

Nisbett RE, Masuda T. Culture and point of view. Proceedings of the National Academy of Sciences (USA) 2003;100:11163–11170.

Ober BA, Dronkers NF, Koss E, Delis DC, Friedland RP. Retrieval from semantic memory in Alzheimer-type dementia. Journal of Clinical and Experimental Neuropsychology 1986;8:75–92. [PubMed: 3944246]

Ostrosky-Solís F, Ardila A, Rosselli M. NEUROPSI: a brief neuropsychological test battery in Spanish with norms by age and educational level. Journal of the International Neuropsychological Society 1999;5:413–433. [PubMed: 10439587]

Park D, Gutchess A. The cognitive neuroscience of aging and culture. Current Directions in Psychological Science 2006;15:105–108.

Park DC, Nisbett R, Hedden T. Aging, culture, and cognition. Journal of Gerontology: Psychological Sciences 1999;54B:P75–P84.

Pekkala S, Albert ML, Spiro A III, Erkinjuntti T. Perseveration in Alzheimer's disease. Dementia and Geriatric Cognitive Disorders 2008;25:109–114. [PubMed: 18075249]

Roberts PM, Le Dorze G. Semantic organization, strategy use, and productivity in bilingual semantic verbal fluency. Brain and Language 1997;59:412–449. [PubMed: 9299071]

Rosen WG. Verbal fluency in aging and dementia. Journal of Clinical Neuropsychology 1980;2:135–146.

Rosselli M, Ardila A. The impact of culture and education on non-verbal neuropsychological measurements: a critical review. Brain and Cognition 2003;52:326–333. [PubMed: 12907177]

Rosselli M, Ardila A, Araujo K, Weekes VA, Caracciolo V, Padilla M, Ostrosky-Solís F. Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. Applied Neuropsychology 2000;7:17–24. [PubMed: 10800624]

Rosselli M, Ardila A, Salvatierra J, Marques M, Matos L, Weekes VA. A cross-linguistic camparison of verbal fluency tests. International Journal of Neuroscience 2002;112:759–776. [PubMed: 12325314]

Rosser A, Hodges JR. Initial letter and semantic category fluency in Alzheimer's disease, Huntington's disease, and progressive supranuclear palsy. Journal of Neurology, Neurosurgery and Psychiatry 1994;57:1389–1394.

Schunn CD, Vera AH. Cross-cultural similarities in category structure. Thinking & Reasoning 2004;10:273–287.

Suhr JA, Jones RD. Letter and semantic fluency in Alzheimer's, Huntington's and Parkinson's dementias. Archives of Clinical Neuropsychology 1998;13:447–454. [PubMed: 14590609]

Traykov L, Baudic S, Raoux N, Latour F, Rieu D, Smagghe A, Rigaud AS. Patterns of memory impairment and perseverative behavior discriminate early Alzheimer's disease from subcortical vascular dementia. Journal of the Neurological Sciences 2005;229–230:75–79.

Troyer AK. Normative data for clustering and switching on verbal fluency task. Journal of Clinical and Experimental Neuropsychology 2000;22:370–378. [PubMed: 10855044]

Troyer AK, Moscovitch M, Winocur G. Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. Neuropsychology 1997;11:138–146. [PubMed: 9055277]

Unverzagt FW, Morgan OS, Thesiger CH, Eldemire DA, Luseko J, Pokuri S, Hui SL, Hall KS, Hendrie HC. Clinical utility of CERAD neuropsychological battery in elderly Jamaicans. Journal of the International Neuropsychological Society 1999;5:255–259. [PubMed: 10217925]

Van der Elst W, Van Boxtel MPJ, Van Breukelen GJP, Jolles J. Normative data for the animal, profession and letter M naming verbal fluency test of Dutch speaking participants and the effects of age, education, and sex. Journal of the International Neuropsychological Society 2006;12:80–89. [PubMed: 16433947]

Vinogradov S, Kirkland J, Poole J, Drexler M, Ober B, Shenaut G. Both processing speed and semantic organization predict verbal fluency in schizophrenia. Schizophrenia Research 2003;59:269–275. [PubMed: 12414084]

Ylikoski, R. The Relationship of Neuropsychological Functioning with Demographic Characteristics, Brain Imaging Findings, and Health in Elderly Individuals. Departments of Psychology and Neurology, University of Helsinki; Helsinki: 2000. p. 379-393.

Yoon C, Feinberg F, Hu P, Gutchess AH, Hedden T, Chen H-YM, Jing Q, Cui Y, Park DC. Category norms as a function of culture and age: comparisons of item responses to 105 categories by American and Chinese adults. Psychology and Aging 2004;19:379–39. [PubMed: 15382989]

Table I

Data derived from different language groups on animal verbal fluency performance (1 minute) in older adults.

Study	Languages (region, country)	n	Age, M (SD)	Education (years) M (SD),	Total number of correct words, M (SD)
Hart, Smith, and Swash (1988)	English (UK)	12	69.8 (6.2)	9.5 (1.0)	8.0 (3.3)
Kontiola et al. (1990)	Finnish (Finland)	86	65.3 (9.1)	6–12 (n/a)	22.4 (5.5)
Binetti et al. (1995)	Italian (Italy)	35	67.4 (7.3)	6.8 (2.7)	14.1 (5.4)
Carew, Lamar, Cloud, Grossman, and Libon (1997)	English (Australia)	31	76.4 (6.6)	12.8 (2.6)	17.7 (4.7)
Troyer et al. (1997)	English (Canada)	54	73.3 (6.5)	13.2 (2.7)	17.8 (4.2)
Kempler et al. (1998)	African Am. Vernacular English	54	72.6 (9.1)	11.6 (4.7)	15.2 (4.4)
	Standard American English	58	76.6 (7.6)	12.3 (3.8)	16.7 (4.2)
	Chinese	67	72.5 (7.3)	10.9 (5.5)	14.3 (5.1)
	Spanish	78	71.9 (7.1)	8.5 (5.4)	12.8 (3.9)
	Vietnamese	60	71.6 (5.8)	8.6 (4.2)	17.3 (5.2)
	(All in California, US)				
Chan and Poon (1999)	Cantonese (Hong Kong)	80	65.7 (3.0)	9.4 (6.1)	15.8 (4.3)
Ostrosky-Solís et al. (1999)	Spanish (Mexico)	50	73.6 (5.4)	7.6 (1.4)	16.6 (4.4)
		50	72.9 (4.8)	13.5 (2.8)	18.4 (4.8)
Unverzagt, Morgan, Thesiger, Eldemire, Luseko, Pokuri, Hui, Hall, and Hendrie (1999)	English (Jamaica)	82	79.3 (6.4)	6.3 (2.8)	9.0 (3.2)
Fillenbaum, Heyman, Huber, Ganguli, and Unverzagt (2001)	African Am. Vernacular English (Indianapolis, US)	83	74.6 (7.1)	10.2 (4.2)	14.4 (n/a)
	African Am. Vernacular English (North Carolina, US)	106	75.9 (13.6)	7.6 (8.7)	12.5 (n/a)
	Standard American English (North Carolina, US)	102	75.8 (19.5)	10.4 (11.9)	14.3 (n/a)
	Standard American English (Monongahela Valley, US)	n/a	73.1 (6.0)	12 (MDN)	14.2 (n/a)
Acevedo et al. (2000)	English	37	50–59 (n/a)	14.4* (2.5)	18.4 (4.9)
		107	60–69 (n/a)		17.1 (4.2)
		172	70–79 (n/a)		15.2 (4.3)
	Spanish (all in Florida, US)	64	50–59 (n/a)	13.4* (3.2)	16.3 (3.9)
		97	60–69 (n/a)		17.2 (5.3)
		76	70–79 (n/a)		16.3 (4.4)

Study	Languages (region, country)	n	Age, M (SD)	Education (years) M (SD),	Total number of correct words, M (SD)
Benito-Cuadrado et al. (2002)	Spanish (Spain)	46	57–65 (n/a)	5–10 (n/a)	17.1 (4.9)
		65	>65 (n/a)	5–10 (n/a)	17.0 (5.8)
		23	58–67 (n/a)	>10 (n/a)	20.1 (6.5)
		32	>67 (n/a)	>10 (n/a)	19.1 (4.7)
Rosselli, Ardila, Araujo, Weekes, Caracciolo, Padilla, and Ostrosky-Solís (2000)	English	45	63.4 (10.1)	16.6 (2.4)	16.8 (5.2)
	Spanish (all in Florida, US)	18	61.3 (8.1)	13.3 (4.8)	16.7 (3.8)
Mathuranath et al. (2003)	Malayalam (India)	50	61.4 (1.4)	7.8 (6.7)	8.3 (3.3)
		85	67.8 (2.7)	7.1 (6.0)	8.0 (3.8)
Kavé (2005)	Hebrew (Israel)	92	59.7 (6.1)	13.9 (2.8)	19.8 (5.1)
		74	76.3 (3.8)	12.5 (2.6)	17.9 (4.7)

Note: The Table is organized chronologically.

*=Based on the mean age and education of all the data introduced in the study. n=number of participants, M=mean, SD=standard deviation, MDN=median, n/a=not available.

Table II

Demographic features of the English-speaking (EN) and the Finnish-speaking (FI) participants.

	EN	FI	χ^{2a}/F*b*	df	p
n	30	30			
Female/male	16/14	14/16	.267[a]	1[a]	.606[a]
Age					
M (SD)	67.5 (5.0)	66.7 (5.5)	.374[b]	1,58[b]	.543[b]
Range	60.0–77.0	60.0–76.0			
Education (years)					
M (SD)	15.0 (2.0)	9.7 (3.3)	55.442[b]	1,58[b]	<.001[b]
Range	12.0–17.0	5.0–17.0			

[a]Chi square test,

[b]analysis of variance (ANOVA). M=mean, SD=standard deviation.

Table III

Data on semantic verbal fluency performance and ANCOVA comparisons between the English-speaking group (EN) and the Finnish-speaking group (FI) co-varying for the educational level of the participants.

Variables	Animals				Clothes				Combined categories			
	EN group (n=30)	FI group (n=30)	ANCOVA F(df=1,57)	p	EN group (n=30)	FI group (n=30)	ANCOVA F(df=1,57)	p	EN group (n=30)	FI group (n=30)	ANCOVA F(df=1,57)	p
Number of correct words												
M (SD)	19.9 (3.8)	18.9 (4.7)	.653	.423	19.0 (4.4)	16.0 (4.5)	.546	.463	39.1 (6.8)	34.9 (8.1)	.000	.990
Range	12.0–26.0	11.0–31.0			12.0–28.9	9.0–24.0			25.0–52.0	20.5–51.0		
Type/token	154/580	133/565	.829[a]	.363[a]	141/574	154/511	2.42[a]	.120[a]	295/1154	287/1076	.208[a]	.648[a]
ratio	=26.5%	=23.5%			=24.6%	=30.1%			=25.6%	=26.7%		
Word length												
M (SD)	1.9 (0.3)	2.5 (0.3)	47.619	<.001	1.6 (0.2)	2.9 (0.3)	219.4	<.001	5.4 (0.5)	3.5 (0.3)	184.150	<.001
range	1.4–2.4	2.1–3.6			1.3–2.0	2.3–3.7			4.5–6.7	2.9–4.3		
Proportion of words in 30 seconds												
M (SD)	66.6 (11.5)	63.6 (10.4)	.407	.526	70.1 (8.7)	70.9 (11.0)	1.148	.289	71.5 (11.0)	70.8 (8.8)	1.127	.293
Range	50.0–100.0	45.0–88.0			51.9–91.7	57.1–100.0			57.6–100.0	52.0–92.5		

M=mean, SD=standard deviation.

[a]=Chi square test (df=1).

Table IV

Frequencies and length of the 10 most common words produced by the English-speaking participants (EN) and the Finnish-speaking participants (FI).

EN participants (n=30)			FI participants (n=30)		
Words	**Frequency and total proportion of words**	**Number of syllables**	**Words (English translation)**	**Frequency and total proportion of words**	**Number of syllables**
Animals					
Dog	29	1	Lehmä (cow)	29	2
Cat	27	1	**Hevonen** (horse)	25	3
Tiger	24	2	**Kissa** (cat)	23	2
Lion	20	2	**Koira** (dog)	23	2
Zebra	18	2	Lammas (sheep)	23	2
Giraffe	18	2	Sika (pig)	20	2
Bear	18	1	**Elefantti** (elephant)	20	4
Elephant	16	3	**Leijona** (lion)	20	3
Horse	16	1	**Karhu** (bear)	19	2
Kangaroo	15	3	**Kirahvi** (giraffe)	14	3
	201/580=34.7%			216/565=38.2%	
Clothes					
Hat	28	1	**Housut** (pants)	29	2
Shirt	23	1	**Sukat** (socks)	26	2
Socks	23	1	**Paita** (shirt)	26	2
Coat	20	1	**Pusero** Blouse/sweater)	24	3
Scarf	20	1	**Takki** (coat)	23	2
Shoes	20	1	**Liivi/liivit** (bra, vest)	22	2
Sweater	20	2	**Kengät** (shoes)	20	2
Gloves	19	1	**Hattu** (hat)	20	2
Bra	18	1	**Käsineet** (gloves)	15	3
Pants	18	1	Hame (skirt)	13	2
	209/574=36.4%			218/511=42.7%	

Note: Words in bold are overlapping between the participant groups.