medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Performance on verbal fluency tasks depends on the given category/letter: Preliminary data from a multivariable analysis

Petar Gabrić[1,2,3,*], ORCiD 0000-0003-2644-194X

Mija Vandek[3], ORCiD 0000-0003-3398-7350

[1]Institute for German Linguistics, Philipps University of Marburg, Marburg, Germany

[2]Department of Psychiatry and Psychotherapy, Philipps University of Marburg, Marburg, Germany

[3]Department of Linguistics, University of Zagreb, Zagreb, Croatia

*Corresponding author: Department of Psychiatry and Psychotherapy, Philipps University of Marburg, Rudolf-Bultmann-Straße 8, 35039 Marburg, Germany; Email: petar.gabric@uni-marburg.de, petar.gabric@protonmail.com

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract: Verbal fluency tasks are often used in neuropsychological research and may have predictive and diagnostic utility in psychiatry and neurology. However, researchers using verbal fluency have uncritically assumed that there are no category- or phoneme-specific effects on verbal fluency performance. We recruited 16 healthy young adult subjects and administered two semantic (animals, trees) and phonemic (K, M) fluency tasks. Because of the small sample size, results should be regarded as preliminary and exploratory. On the animal compared to the tree task, subjects produced significantly more legal words, had a significantly lower intrusion rate, significantly shorter first-response latencies and final silence periods, as well as significantly shorter between-cluster response latencies. These differences may be explained by differences in the category sizes, integrity of the categories' borders, and efficiency of the functional connectivity between subcategories. On the K compared to the M task, subjects produced significantly more legal words and had significantly shorter between-cluster response times. Counterintuitively, a corpus analysis revealed there are more words starting with ⟨m⟩ compared to ⟨k⟩ in the experimental language. Our results potentially have important implications for research utilizing verbal fluency, including decreased reproducibility, questionable reliability of diagnostic and predictive tools based on verbal fluency, decreased knowledge accumulation, and increased number of publications with potentially misleading clinical interpretations.

Keywords: semantic fluency; phonemic fluency; neuropsychological assessment; behavioral neuroscience; semantic memory

1. INTRODUCTION

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

1.1. Verbal fluency in basic and clinical neuroscience

Verbal fluency (VF) is a common neuropsychological tool used for assessing cognitive and linguistic abilities in both healthy and clinical populations (Oberg & Ramírez, 2006; Pekkala, 2012; Stielow, 2017; Thiele et al., 2016; Turkstra, 2018). Most commonly, the semantic (SF) and phonemic fluency (PF) tasks are administered, in which the subject is to name as many words of the given semantic category (e.g., *animals*) or beginning with the given letter/phoneme (e.g., *F*), respectively, in limited time. Usually, only the productivity score (i.e., the raw number of legal words) is calculated.

The relative pervasiveness of SF and PF tasks in neuropsychological research can at least partly be attributed to their simplicity and brevity which increase the probability that the subject will understand the task and maintain the instructions in their working memory throughout the task. This makes VF particularly suitable for use in psychiatric and neurological populations where the presence of psychopathological phenomena might hinder understanding of the task and its completion (Holmlund et al., 2019). Furthermore, the task requires continuous behavioral responses from the subject, making it appropriate for studying brain metabolism during cognitive tasks via neuroimaging techniques (Costafreda et al., 2006; Gonzalez-Burgos et al., 2021; Kircher et al., 2011; Nagels et al., 2011; Scheuringer et al., 2020; Vonk et al., 2019; Wende et al., 2012). Both SF and PF have been used for studying functional brain metabolism in diverse psychiatric disorders, including depression, schizophrenia, bipolar disorder, etc. (Backes et al., 2014; Corrigendum, 2015; Erratum, 2021; Husain et al., 2020; Takizawa et al., 2014; Tassi et al., 2022; Yeung et al., 2021), as well as in studies where psychiatric symptoms are induced in healthy subjects, e.g., via infusion of subanesthetic ketamine (Nagels et al., 2011a, b).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

VF tasks are further used for establishing the presence of specific cognitive deficits in pathological populations (Stielow & Stenneken, 2017). Deficits in the productivity on SF and LF in relation to healthy subjects have been used as arguments for deficits in semantic memory, executive functioning, linguistic processing, etc. (Aita et al., 2019; Amunts et al., 2020, 2021; Kavé et al., 2020; Ledoux et al., 2014; Shao et al., 2014; Ventura et al., 2005; Whitesite et al., 2016) in pathological populations such as patients with mild cognitive impairment (Bauer et al., 2021), Alzheimer's disease (Henry et al., 2004), Parkinson's disease (Henry & Crawford, 2004c; Højlund et al., 2017), Huntington's disease (Henry et al., 2005), traumatic brain injury (Cermak et al., 2021; Henry & Crawford, 2004a, b; Rosenkranz et al., 2020; Thiele et al., 2016), multiple sclerosis (Santangelo et al., 2019), epilepsy (Metternich et al., 2014), schizophrenia (Bokat & Goldberg, 2003; Henry & Crawford, 2005a; Rosenkranz et al., 2019; Tan et al. 2020, 2021; cf. Gabrić, 2021a, b), depression (Henry & Crawford, 2005b; Klumpp & Deldin, 2010), bipolar disorder (Raucher-Chéné et al., 2017), obsessive-compulsive disorder (Henry, 2006), autism (Spek et al., 2009), specific language impairment and other developmental disorders (Mengisidou, 2019), HIV infection (Iudicello et al., 2006), and many others.

Furthermore, there is an increasing number of studies suggesting that different behavioral and neural measurements of both SF and PF may diagnostically specific for different types of pathologies and/or discriminate between healthy individuals and patients of different pathologies, including cognitive impairment and dementia (Jacobs et al., 2020; Lam et al., 2006; Linz et al., 2017; McDonnell et al., 2019; Radanovic et al., 2007, 2009), schizophrenia (Becker et al., 2009; Costafreda et al., 2011; Takizawa et al., 2014), bipolar disorder (Costafreda et al., 2011), localization of cerebral lesions (Pendleton et al., 1982; cf. Turkstra et al., 2005; Wong et al., 2010), and others. VF may be used to discriminate between mild cognitive impairment and

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

dementia (Bertola et al., 2014; Zhao et al., 2013), between vascular cognitive impairment without dementia and dementia (Zhao et al., 2013), and between idiopathic normal pressure hydrocephalus and dementia (Louta et al., 2017). Several studies have shown that VF tasks may be used as tools for early prediction of cognitive impairment and dementia (Clark et al., 2016; Cosentino et al., 2006; Gomez et al., 2006; Pakhomov et al., 2018; Sutin et al., 2019; Tröger et al., 2018; Vaughan et al., 2018).

VF tasks are also used within language lateralization assessments (Birn et al. 2010; Bishop et al. 2009; Knecht et al. 2000; Meyer et al. 2014; Uomini et al., 2013) and they are used as a measure of verbal intelligence (Crawford et al., 1992; Miller, 1984).

1.2. Category-specific and letter-specific effects on verbal fluency performance

Most researchers utilizing SF and PF uncritically assume that there are no category-specific or phoneme/letter-specific effects on the quantity and quality of performance on different SF and PF tasks. This is evidenced (1) from the fact that different studies draw the same inferences about cognitive deficits or other phenomena using different semantic categories and/or letters and (2) from the fact that studies administering multiple semantic categories and/or letters typically calculate only the aggregate score for the different SF and/or PF tasks. Despite this silent and uncritical assumption present in much of the literature, there are no data to support this assumption and, more problematically, there are clear indications in the linguistic and neuropsychological literature that differences in performance patterns on different semantic categories and letters should be expected. Given that different performance patterns might reflect different underlying cognitive processes (Vandek et al., 2018) and that VF tasks are often used for exploring cognitive processes in pathological populations, the silent assumption of SF and PF

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

homogeneity might have important implications for research using SF and PF tasks and might explain some reproducibility issues in research using VF.

Neuroscientific research has clearly demonstrated that semantic memory is widely distributed over the cortical surface and the processing of different semantic categories may be associated with disproportionate levels of metabolic activity in different brain regions (Dreyer & Pulvermüller, 2018; Fahimi Hnazaee et al., 2018; Gabrić, 2019, 2021c; Gainotti, 2015; Huth et al., 2016; Liuzzi et al., 2020). Furthermore, both theoretical and experimental linguistics agree that different lexical-semantic features such as concreteness/abstractness and animacy/inanimacy underlie divergent cognitive processing mechanisms, while pathological populations may exhibit feature-specific deficits on lexical tasks (Carammaza & Shelton, 1998; de Almeida et al., 2021; De Letter et al., 2019; Seijdel et al., 2021; Sekulić Sović et al., 2019). Similarly, processing of different phonological categories is associated with different cognitive and neural mechanisms (Kronrod et al., 2016; Lago et al., 2015).

Despite this, the effect of the given category or phoneme/letter on performance on SF and PF has been alarmingly understudied. It is clear from studies reporting descriptive data of multiple SF tasks that there are stark discrepancies in the number of legal words produced by healthy subjects across different semantic categories. Typically, adult healthy subjects produce the most words on animal fluency, while productivity on other semantic categories may be reduced by more than twofold (Bosanac et al., 2017; Gabrić & Vandek, 2021a; Jebahi et al. 2020, 2021; Kim et al., 2015; Moreno-Martínez et al., 2017; Vicente et al., 2021). For example, Jebahi et al. (2020) reported that subjects produced 18 words on average on animal fluency but less than ten on the categories *accessories*, *tools*, *transportation means*, *furniture*, and *sports*. One study found that performance on categories *animals* and *vegetables* was associated with neither age nor

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

education, unlike *fruits* (Farghaly et al., 2018). Moreover, pathological populations may show deficits of different degrees on different SF categories or they may show normative deficits on one but not the other category. One study on SF in Alzheimer's disease has revealed that the extent of the deficit is dependent on the given category's size (Diaz et al., 2004), while Moreno-Martínez et al. (2017) and Neves et al. (2020) found that although the animal fluency is by far the most commonly utilized SF task for research on dementia, other categories such as *supermarket items*, *fruits*, and *vegetables* are more sensitive in distinguishing healthy people from patients with dementia. Gabrić and Vandek (2021a) reported that inpatients with first-episode psychosis displayed deficient clustering on the categories *animals*, *vegetables*, and *musical instruments* but not on *trees* and *fruits*, despite the presence of a normative deficit in productivity across all categories, while Tagini et al. (2021) reported that patients with Parkinson's disease showed semantically disorganized output on the category *fruits* but not *animals*.

Similarly, there is a long-standing division between "easy" and "difficult" letters in English within PF assessments (Borkowski et al. 1967), recently supported by Barry et al. (2008) who reported from a meta-analysis that performance on the CFL version (i.e., letters *C*, *F*, and *L*) of the Controlled Oral Word Association Test (COWAT) was significantly reduced and thus more difficult compared to the FAS version, as well as that healthy subjects exhibit greater variability on the FAS compared to the CFL version (cf. Ferrett et al., 2014).

The eventual existence of category- and phoneme/letter-specific effects on SF and PF performances may have important implications for the past and current research using VF, including decreased reproducibility, questionable reliability of diagnostic and predictive tools based on verbal fluency, decreased knowledge accumulation, and increased number of

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

publications with potentially misleading clinical interpretations. Although there exist normative values for VF performance which are category- or letter-specific, they are not always used for interpretation in psychiatric and neurological research and in almost all of the cases, they relate to the raw number of legal words the subject produced, ignoring thus a number of other more specific aspects of VF performance.

The aim of the present study was to investigate whether there are differences in performance patterns within different SF and LF tasks. We compared performances in multiple variables between two SF and two LF tasks.

2. MATERIALS AND METHODS

2.1. Subjects

Sixteen healthy subjects were recruited for the study (6 males, mean age 22.250 ± 2.543 years, mean education 15.188 ± 1.759 years). All were native speakers of one of the Croatian dialects, students at the University of Zagreb, young adults, and right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Subjects who had more than one native language or had a history of psychiatric or neurological disorders were not included in the study. Ethical approval was obtained by the Ethical Committee of the Faculty of Humanities and Social Sciences, University of Zagreb (dated 2018/10/03). All subjects signed an informed consent form.

2.2. Verbal fluency assessment

Subjects were to name as many words according to the given cue in 60 seconds. The categories *animals* and *trees* were used for SF, while the letters ⟨k⟩ and ⟨m⟩ (corresponding to phonemes /k/

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

and /m/ in Croatian dialects) were used for PF. An exploration of the Croatian web corpus hrWaC (Ljubešić & Erjavec, 2015) revealed that there were more lemmas starting with ⟨m⟩ (N = 436 719) compared to ⟨k⟩ (N = 412 961) (Goodness of fit: $\chi^2_{(1)} = 664.300$, $p < .00001$). PF was administered one to two weeks after SF to avoid SF-related recency effects on PF. Responses were audiotaped and transcribed using ELAN (Version 5.8) (ELAN, 2019; Wittenburg et al., 2006). Clusters were identified according to Troyer et al. (1997), with the cluster threshold set to 2 words (Abwender et al., 2001; Lehtinen et al., 2021; Troyer & Moscovitch, 2006). Two native speakers independently rated the outputs. Discrepancies were discussed until mutual agreement was achieved. Clustering and switching variables were calculated according to Gabrić and Vandek (2021a).

Dependent variables included: 1) legal words (raw), 2) perseveration rate (%), 3) intrusion rate (%), 4) first-response latency (ms), 5) final silence (ms), 6) cluster magnitude (words per cluster), 7) switching rate (%), 8) between-cluster response latencies (ms), and 9) within-cluster response latencies (ms).

2.4. Statistics

Statistical analyses were performed in JASP (Version 0.11.1.0) (JASP Team, 2020). Comparisons between the animal and tree variables, and K and M variables were performed using the paired-sample Wilcoxon signed-rank test. The intrusion rates on SF were compared with the one-sample Wilcoxon signed-rank test with the test value set at 1 ± 1, as no subjects produced intrusions on the animal task. The effect sizes are given by the matched rank biserial correlation. Correlations were explored using Spearman correlation coefficients.

3. RESULTS

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

3.1. Comparisons and correlations within semantic fluency tasks

Results from the comparisons are displayed in Table 1. On the animal compared to the tree task, subjects produced significantly more legal words, had a significantly lower intrusion rate, significantly shorter first-response latencies and final silence periods, as well as significantly shorter between-cluster response latencies. All of the effect sizes but the one for final silence were strong. Additionally, subjects displayed significantly shorter within-cluster compared between-cluster response latencies on both the animal and tree tasks ($Z = 119.0$, $p < .001$, $r_b = 0.983$ and $Z = 120.0$, $p < .001$, $r_b > .999$, respectively).

Table 1. Within-sample comparisons between the animal and tree tasks

	Animals	Trees	Statistics
Legal words (raw)	23.667 (6.114)	9.400 (2.293)	$Z = 120.0$, $p < .001$, $r_b > .999$
Perseveration rate (%)	1.0 (2.6)	0.6 (2.2)	$Z = 3.0$, $p = .584$, $r_b = .400$
Intrusion rate (%)*	0.0 (0.0)	4.7 (7.1)	$Z = 21.0$, $p = .036$, $r_b > .999$
First-response latency (ms)	1029.333 (360.723)	1789.333 (1064.341)	$Z = 11.0$, $p = .006$, $r_b = .816$
Final silence (ms)	5015.067 (3883.662)	14217.067 (12080.566)	$Z = 22.0$, $p = .030$, $r_b = .633$
Cluster magnitude (words)	3.331 (0.776)	2.961 (0.649)	$Z = 81.0$, $p = .244$, $r_b = .350$
Switching rate (%)	31.3	43.5	$Z = 22.0$, $p = .108$, $r_b = .516$

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

	(15.0)	(14.6)	
Between-cluster response latencies (ms)	1706.500 (686.115)	5979.967 (3234.909)	$Z = 119.0$, $p < .001$, $r_b = .983$
Within-cluster response latencies (ms)	427.333 (278.240)	657.833 (563.643)	$Z = 79.0$, $p = .293$, $r_b = .317$

Note: Mean values are reported. Standard deviations appear in parentheses. *one-sample Wilcoxon-signed rank test

The number of legal words on both the animal and tree tasks was significantly positively correlated with cluster magnitude ($\rho = .547$, $p = .035$ and $\rho = .753$, $p = .001$, respectively). No other correlations were significant.

3.2. Comparisons and correlations within phonemic fluency tasks

Results from the comparisons are displayed in Table 2. On the K compared to the M task, subjects produced significantly more legal words and had significantly shorter between-cluster response latencies. Additionally, there were no significant differences between the between-cluster and within-cluster response latencies on either the K or M task ($Z = 31.0$, $p = .569$, $r_b = .205$ and $Z = 25.0$, $p = .301$, $r_b = .359$, respectively).

Table 2. Within-sample comparisons between the K and M tasks

	K	M	Statistics
Legal words (raw)	15.750 (2.864)	12.833 (4.086)	$Z = 58.0$, $p = .029$, $r_b = .758$
Perseveration rate (%)	0.4 (1.3)	1.2 (2.9)	$Z = 0.0$, $p = .371$, $r_b > .999$

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Intrusion rate (%)	0.0 (0.0)	0.0 (0.0)	N/A
First-response latency (ms)	2409.167 (1709.005)	1893.333 (640.289)	Z = 50.0, p = .424, r_b = .282
Final silence (ms)	3740.500 (3270.119)	6285.333 (6335.809)	Z = 30.0, p = .519, r_b = .231
Cluster magnitude (words)	2.562 (0.770)	2.540 (0.830)	Z = 25.0, p = .813, r_b = .111
Switching rate (%)	72.3 (13.9)	71.4 (11.4)	Z = 29.000, p = .756, r_b = .121
Between-cluster response latencies (ms)	2014.583 (787.570)	2732.708 (1158.775)	Z = 9.0, p = .016, r_b = .769
Within-cluster response latencies (ms)	1864.375 (1347.603)	2169.375 (1698.251)	Z = 39.0, p > .999, r_b < .001

Note: Mean values are reported. Standard deviations appear in parentheses.

The number of legal words on the K task was significantly positively and strongly correlated with cluster magnitude (ρ = .731, p = .007), while on the M task none of the correlations were significant. No other correlations were significant.

3.3. Correlations with sociodemographic variables

As expected in a sample of university students, age and education were significantly positively and strongly correlated (ρ = .746, p < .001).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Both age and education were significantly positively and moderately correlated with cluster magnitude on the animal task (ρ = .659, p = .008 and ρ = .689, p = .005, respectively) but there were no significant correlations between the tree task and sociodemographic variables.

Age was significantly negatively and moderately correlated with the first-response latency on the K task (ρ = -.623, p = .030). There were no significant correlations between the M task and sociodemographic variables.

4. DISCUSSION

4.1. Semantic fluency

Comparisons between the animal and tree fluency tasks revealed highly different performance patterns across the two tasks. Subjects produced significantly 60.28 % more legal words on the animal compared to the tree category, suggesting that the subjects in our sample had access to more animal than tree words, as well as that the animal task was appreciably less difficult compared to the tree task. Both the former and the latter suggestions are also arguably indicated by the significant differences in the duration of the final silence between the two tasks. On average, subjects stopped responding in the 55. second on the animal task compared to the 46. second on the tree task, with the latter possibly indicating that the readily available tree words were more or less exhausted before the end of the task and that the access to less familiar (but entrenched) words was much more demanding in the tree compared to the animal task. Subjects were further significantly slower to utter the first response on the tree compared to the animal task, indicating that it was more difficult for the subjects to access the category *trees*, retrieve, and utter the first tree word compared to the category *animals*. The results regarding first-

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

response latency and final silence indicate that the tree task was more difficult for the subjects compared to the animal task during both the incipient and ending stages of the task.

Significant differences between the two tasks were identified in the intrusion rate as well. While no subjects produced intrusions on the animal task, intrusions on the tree task were produced by seven subjects. We would like to note at this point that we were very careful in classifying responses as intrusions and that our criteria for intrusion classification were more lax (but nevertheless precise) compared to, for example, Moreno-Martínez et al. (2017) who classified fruit trees (such as *pear*) as intrusions on the tree task. In our study, fruit trees were classified as legal because we found there was no reason to classify them as errors. Intrusions on the tree task included: *glog* 'hawthorn' (2), *bugenvilija* 'bougainvillea' (1), *hortenzija* 'hortensia' (1), *ebanovina* 'ebony' (1), *ruža* 'rose' (1), *jorgovan* 'lilac' (1), *loza* 'grapevine', and *žir* 'acorn' (1). The significantly higher intrusion rate on the tree compared to the animal task indicates that the boundaries of the category *trees* were considerably looser compared to the boundaries of the category *animals* in our sample.

Subjects did not exhibit significant differences in the clustering and switching strategies between the two SF tasks. However, while descriptive data showed only slight differences in cluster magnitude, the differences in the switching rate were more pronounced, suggesting that subjects were less likely to produce a cluster once a switch has been uttered on the tree compared to the animal task. This might reflect lesser functional connectivity between concepts in the category *trees* compared to *animals* and, specifically, lesser hypernymic/hyponymic organization within the category *trees* compared to *animals*. In any case, the existence of clusters on the tree task still indicates the presence of hypernymic/hyponymic and co-hyponymic structures in the category *trees*. Subjects were significantly slower in between-cluster response latencies on the tree

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

compared to the animal task indicating that the functional connectivity between subcategories is less efficient in the category *trees* compared to *animals*. Interestingly, larger clusters were associated with better productivity on both the animal and tree tasks, yet the correlation coefficient was considerably higher in the tree task. Thus, those who had better organization of hypernymic/hyponymic and co-hyponymic relations between concepts in both categories produced more words on the given SF task and this effect was more pronounced on the animal task.

4.2. Phonemic fluency

The differences in performance patterns between the two PF tasks were much less pronounced than between the two SF tasks. Significant differences were observed in the productivity and between-cluster response latencies, with subjects producing significantly more words and having significantly shorter between-cluster response latencies on the K compared to the M task. These results suggest that performance on the K task was easier for the subjects compared to the M task. Because there are presumably more words starting with ⟨m⟩ compared to the ⟨k⟩ task, this result presumably does not reflect the disproportionate number of words starting with either ⟨k⟩ or ⟨m⟩ in the Croatian dialects.

5. CONCLUSION

Although preliminary, our results potentially have important implications for neuropsychological research utilizing VF tasks. VF studies typically draw the same inferences from different categories or letters and in the case of using multiple categories or letters, the aggregate result of SF and LF is typically calculated and used in analyses. Our results show that performance may vary depending on the given category or letter and we argue that this may reflect differential

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

linguistic processes. We further find it likely that these disproportionalities also reflect differential cognitive processes, as those assessed by standard neuropsychological tasks used in psychiatric and neurological research. Thus, we do not agree that the same inferences can be drawn from different categories or letters. Given the frequent use of VF in psychiatry, neurology, and other disciplines, there is an urgent need to reexamine our understanding of VF and its associations with cognitive and neural processes, e.g., by investigating behavioral, cognitive, and neural correlates of different categories and letters and subsequently to revisit previous research.

Acknowledgments

The content of the manuscript has previously appeared online as a preprint (Gabrić & Vandek, 2021b, 2022). Parts of this research were presented on October 23rd 2020 at the 12th Annual Meeting of the Society for the Neurobiology of Language (Gabrić & Vandek, 2020).

Funding

Open Access funding was enabled by the Publications Fund of the University Library of the Philipps University of Marburg. This research received no specific grant from any funding agency, commercial or nonprofit sectors.

Ethics statement

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Ethical approval was obtained by the Ethical Committee of the Faculty of Humanities and Social Sciences, University of Zagreb (dated 2018/10/03). All subjects signed an informed consent form.

Conflict of Interest

Authors state no conflict of interest.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*(3), 323–338. https://doi.org/10.1177/107319110100800308

Aita, S. L., Beach, J. D., Taylor, S. E., Borgogna, N. C., Harrell, M. N., & Hill, B. D. (2019). Executive, language, or both? An examination of the construct validity of verbal fluency measures. *Applied Neuropsychology: Adult, 26*(5), 441–451. https://doi.org/10.1080/23279095.2018.1439830

Amunts, J., Camilleri, J. A., Eickhoff, S. B., Heim, S., & Weis, S. (2020). Executive functions predict verbal fluency scores in healthy participants. *Scientific Reports, 10*(1), 11141. https://doi.org/10.1038/s41598-020-65525-9

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Amunts, J., Camilleri, J. A., Eickhoff, S. B., Patil, K. R., Heim, S., von Polier, G. G., & Weis, S. (2021). Comprehensive verbal fluency features predict executive function performance. *Scientific Reports*, *11*(1), 6929. https://doi.org/10.1038/s41598-021-85981-1

Backes, H., Dietsche, B., Nagels, A., Stratmann, M., Konrad, C., Kircher, T., & Krug, A. (2014). Increased neural activity during overt and continuous semantic verbal fluency in major depression: Mainly a failure to deactivate. *European Archives of Psychiatry and Clinical Neuroscience*, *264*(7), 631–645. https://doi.org/10.1007/s00406-014-0491-y

Barry, D., Bates, M. E., & Labouvie, E. (2008). FAS and CFL forms of verbal fluency differ in difficulty: A meta-analytic study. *Applied Neuropsychology*, *15*(2), 97–106. https://doi.org/10.1080/09084280802083863

Bauer, K., & Malek-Ahmadi, M. (2021). Meta-analysis of Controlled Oral Word Association Test (COWAT) FAS performance in amnestic mild cognitive impairment and cognitively unimpaired older adults. *Applied Neuropsychology: Adult*, 1–7 [Online ahead of print]. https://doi.org/10.1080/23279095.2021.1952590

Becker, H. E., Nieman, D. H., Dingemans, P. M., van de Fliert, J. R., De Haan, L., & Linszen, D. H. (2010). Verbal fluency as a possible predictor for psychosis. *European Psychiatry*, *25*(2), 105–110. https://doi.org/10.1016/j.eurpsy.2009.08.003

Bertola, L., Mota, N. B., Copelli, M., Rivero, T., Diniz, B. S., Romano-Silva, M. A., Ribeiro, S., & Malloy-Diniz, L. F. (2014). Graph analysis of verbal fluency test discriminate between patients with Alzheimer's disease, mild cognitive impairment and normal elderly controls. *Frontiers in Aging Neuroscience*, *6*, 185. https://doi.org/10.3389/fnagi.2014.00185

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Birn, R. M., Kenworthy, L., Case, L., Caravella, R., Jones, T. B., Bandettini, P. A., & Martin, A. (2010). Neural systems supporting lexical search guided by letter and semantic category cues: A self-paced overt response fMRI study of verbal fluency. *NeuroImage, 49*(1), 1099–1107. https://doi.org/10.1016/j.neuroimage.2009.07.036

Bishop, D. V. M., Watt, H., & Papadatou-Pastou, M. (2009). An efficient and reliable method for measuring cerebral lateralization during speech with functional transcranial Doppler ultrasound. *Neuropsychologia, 47*(2), 587–590. https://doi.org/10.1016/j.neuropsychologia.2008.09.013

Bokat, C. E., & Goldberg, T. E. (2003). Letter and category fluency in schizophrenic patients: A meta-analysis. *Schizophrenia Research, 64*(1), 73–78. https://doi.org/10.1016/S0920-9964(02)00282-7

Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia, 5*(2), 135–140. https://doi.org/10.1016/0028-3932(67)90015-2

Bosanac, B., Dominik, P., Gabrić, P., Hrstić, A., Kelčec Ključarić, D., Kužina, I., Sekulić Sović, M., & Vandek, M. (2017). Lexical-semantic categorisation in patients of schizophrenia [Conference poster]. *4th Zadar Linguistic Forum "Research Methods in Applied Linguistics"*, Department of Linguistics, University of Zadar, Zadar, Croatia. https://doi.org/10.5281/zenodo.4681378

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience, 10*(1), 1–34. https://doi.org/10.1162/089892998563752

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Cermak, C. A., Scratch, S. E., Kakonge, L., & Beal, D. S. (2021). The effect of childhood traumatic brain injury on verbal fluency performance: A systematic review and meta-analysis. *Neuropsychology Review*, *31*(1), 1–13. https://doi.org/10.1007/s11065-020-09475-z

Clark, D. G., McLaughlin, P. M., Woo, E., Hwang, K., Hurtz, S., Ramirez, L., … Apostolova, L. G. (2016). Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *2*(1), 113–122. https://doi.org/10.1016/j.dadm.2016.02.001

Corrigendum to "Neuroimaging-aided differential diagnosis of the depressive state" [Neuroimage 85: 498–507, 2014]. (2015). *NeuroImage*, *109*, 530. https://doi.org/10.1016/j.neuroimage.2015.01.021

Cosentino, S., Scarmeas, N., Albert, S. M., & Stern, Y. (2006). Verbal fluency predicts mortality in Alzheimer disease. *Cognitive and Behavioral Neurology*, *19*(3), 123–129. https://doi.org/10.1097/01.wnn.0000213912.87642.3d

Costafreda, S. G., Fu, C. H. Y., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus. *Human Brain Mapping*, *27*(10), 799–810. https://doi.org/10.1002/hbm.20221

Costafreda, S. G., Fu, C. H., Picchioni, M., Toulopoulou, T., McDonald, C., Kravariti, E., … McGuire, P. K. (2011). Pattern of neural responses to verbal fluency shows diagnostic specificity for schizophrenia and bipolar disorder. *BMC Psychiatry*, *11*(1), 18. https://doi.org/10.1186/1471-244x-11-18

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Crawford, J. R., Moore, J. W., & Cameron, I. M. (1992). Verbal fluency: A NART based equation for the estimation of premorbid performance. *British Journal of Clinical Psychology, 31*(3), 327–329. https://doi.org/10.1111/j.2044-8260.1992.tb00999.x

de Almeida, R. G., Mobayyen, F., Antal, C., Kehayia, E., Nair, V. P., & Schwartz, G. (2021). Category-specific verb-semantic deficits in Alzheimer's disease: Evidence from static and dynamic action naming. *Cognitive Neuropsychology, 38*(1), 1–26. https://doi.org/10.1080/02643294.2020.1858772

De Letter, M., Van Borsel, J., Lanckmans, E., Batens, K., Hemelsoet, D., Duyck, W., Fias, W., & Santens, P. (2019). Category specific recall in acute stroke: A case with letter speech. *Neurocase, 25*(6), 251–258. https://doi.org/10.1080/13554794.2019.1673440

Diaz, M., Sailor, K., Cheung, D., & Kuslansky, G. (2004). Category size effects in semantic and letter fluency in Alzheimer's patients. *Brain and Language, 89*(1), 108–14. https://doi.org/10.1016/S0093-934X(03)00307-9

Dreyer, F. R., & Pulvermüller, F. (2018). Abstract semantics in the motor system? An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning. *Cortex, 100*, 52–70. https://doi.org/10.1016/j.cortex.2017.10.021

ELAN (Version 5.8) [Computer software]. (2019). Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen. https://archive.mpi.nl/tla/elan

Erratum regarding previously published research papers. (2021). *EBioMedicine, 66*, 103295. https://doi.org/10.1016/j.ebiom.2021.103295

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Fahimi Hnazaee, M., Khachatryan, E., & Van Hulle, M. M. (2018). Semantic features reveal different networks during word processing: An EEG source localization study. *Frontiers in Human Neuroscience*, *12*, 503. https://doi.org/10.3389/fnhum.2018.00503

Farghaly, M., Hussein, M., Hassan, A., Hegazy, M., & Sabbah, A. (2018). Testing of verbal fluency in Egyptians: Cultural and educational challenges. *Cognitive and Behavioral Neurology*, *31*(3), 133–141. https://doi.org/10.1097/wnn.0000000000000160

Ferrett, H. L., Carey, P. D., Baufeldt, A. L., Cuzen, N. L., Conradie, S., Dowling, T., … Thomas, K. G. F. (2014). Assessing phonemic fluency in multilingual contexts: Letter selection methodology and demographically stratified norms for three South African language groups. *International Journal of Testing*, *14*(2), 143–167. https://doi.org/10.1080/15305058.2013.865623

Gabrić, P. (2019). *Evolucija sintakse i semantike iz psiholingvističke i neurolingvističke perspektive*. M.A. thesis: University of Zagreb, Zagreb, Croatia. https://doi.org/10.31237/osf.io/3ap9u

Gabrić, P. (2021a). Comment on "Semantic memory impairment across the schizophrenia continuum: A meta-analysis of category fluency performance" by Tan et al. [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/jcnpt

Gabrić, P. (2021b). Comment on "Corrigendum to: Semantic memory impairment across the schizophrenia continuum: A meta-analysis of category fluency performance" by Tan et al. [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/ayfb7

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Gabrić, P. (2021c). *Evolution of Syntax and Semantics from Psycholinguistic and Neurolinguistic Perspectives*. M.A. thesis [English translation]: University of Zagreb, Zagreb, Croatia. https://doi.org/10.31234/osf.io/e9w4u

Gabrić, P., & Vandek, M. (2020, October 23). Quantitative and qualitative differences in performance within the semantic and letter fluency tasks [Conference poster]. *12th Annual Meeting of the Society for the Neurobiology of Language*, online, US. https://doi.org/10.5281/zenodo.4681799

Gabrić, P., & Vandek, M. (2021a). Semantic fluency reveals reduced functional connectivity between subcategorical co-hyponyms in recent-onset inpatients with first-episode psychosis. *Clinical Linguistics & Phonetics*, 1–17 [Online ahead of print]. https://doi.org/10.1080/02699206.2021.1961019

Gabrić, P., & Vandek, M. (2021b). Performance on verbal fluency tasks depends on the given category/letter: Preliminary data from a multivariable analysis [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/jykpu

Gabrić, P., & Vandek, M. (2022). Performance on verbal fluency tasks depends on the given category/letter: Preliminary data from a multivariable analysis [Preprint]. *medRxiv*. https://doi.org/10.1101/2021.12.30.21268567

Gainotti, G. (2015). Inborn and experience-dependent models of categorical brain organization: A position paper. *Frontiers in Human Neuroscience*, *9*, 2. https://doi.org/10.3389/fnhum.2015.00002

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Gomez, R. G., & White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology, 21*(89), 771–775. https://doi.org/10.1016/j.acn.2006.06.012

Gonzalez-Burgos, L., Pereira, J. B., Mohanty, R., Barroso, J., Westman, E., & Ferreira, D. (2021). Cortical networks underpinning compensation of verbal fluency in normal aging. *Cerebral Cortex, 31*(8), 3832–3845. https://doi.org/10.1093/cercor/bhab052

Henry, J. (2006). A meta-analytic review of Wisconsin Card Sorting Test and verbal fluency performance in obsessive-compulsive disorder. *Cognitive Neuropsychiatry, 11*(2), 156–176. https://doi.org/10.1080/13546800444000227

Henry, J. D., & Crawford, J. R. (2004a). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18*(2), 284–295. https://doi.org/10.1037/0894-4105.18.2.284

Henry, J. D., & Crawford, J. R. (2004b). A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. *Neuropsychology, 18*(4), 621–628. ttps://doi.org/10.1037/0894-4105.18.4.621

Henry, J. D., & Crawford, J. R. (2004c). Verbal fluency deficits in Parkinson's disease: A meta-analysis. *Journal of the International Neuropsychological Society, 10*(4), 608–622. https://doi.org/10.1017/s1355617704104141

Henry, J., & Crawford, J. (2005a). A meta-analytic review of verbal fluency deficits in schizophrenia relative to other neurocognitive deficits. *Cognitive Neuropsychiatry, 10*(1), 1–33. https://doi.org/10.1080/13546800344000309

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Henry, J. D., & Crawford, J. R. (2005b). A meta-analytic review of verbal fluency deficits in depression. *Journal of Clinical and Experimental Neuropsychology*, *27*(1), 78–101. https://doi.org/10.1080/138033990513654

Henry, J. D., & Beatty, W. W. (2006). Verbal fluency deficits in multiple sclerosis. *Neuropsychologia*, *44*(7), 1166–1174. https://doi.org/10.1016/j.neuropsychologia.2005.10.006

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, *42*(9), 1212–1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2005). A meta-analytic review of verbal fluency deficits in Huntington's disease. *Neuropsychology*, *19*(2), 243–252. https://doi.org/10.1037/0894-4105.19.2.243

Herrera-García, J. D., Rego-García, I., Guillén-Martínez, V., Carrasco-García, M., Valderrama-Martín, C., Vílchez-Carrillo, R., … Carnero-Pardo, C. (2019). Discriminative validity of an abbreviated Semantic Verbal Fluency Test. *Dementia & Neuropsychologia*, *13*(2), 203–209. https://doi.org/10.1590/1980-57642018dn13-020009

Højlund, A., Petersen, M. V., Sridharan, K. S., & Østergaard, K. (2017). Worsening of verbal fluency after deep brain stimulation in Parkinson's disease: A focused review. *Computational and Structural Biotechnology Journal*, *15*, 68–74. https://doi.org/10.1016/j.csbj.2016.11.003

Holmlund, T. B., Cheng, J., Foltz, P. W., Cohen, A. S., & Elvevåg, B. (2019). Updating verbal fluency analysis for the 21st century: Applications for psychiatry. *Psychiatry Research*, *273*, 767–769. https://doi.org/10.1016/j.psychres.2019.02.014

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Husain, S. F., Tang, T.-B., Yu, R., Tam, W. W., Tran, B., Quek, T. T., … Ho, R. C. (2020). Cortical haemodynamic response measured by functional near infrared spectroscopy during a verbal fluency task in patients with major depression and borderline personality disorder. *EBioMedicine*, *51*, 102586. https://doi.org/10.1016/j.ebiom.2019.11.047

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458. https://doi.org/10.1038/nature17637

Iudicello, J. E., Woods, S. P., Parsons, T. D., Moran, L. M., Carey, C. L., & Grant, I. (2006). Verbal fluency in HIV infection: A meta-analytic review. *Journal of the International Neuropsychological Society*, *13*(1), 183–189. https://doi.org/10.1017/s1355617707070221

Jacobs, S., Mercuri, G., & Holtzer, R. (2020). Assessing within-task verbal fluency performance: The utility of individual time intervals in predicting incident mild cognitive impairment. *Aging, Neuropsychology, and Cognition*, *28*(5), 733–747. https://doi.org/10.1080/13825585.2020.1817306

JASP Team. (2020). *JASP* (Version 0.13.1) [Computer software]. University of Amsterdam. https://jasp-stats.org

Jebahi, F., Jaoude, R. A., & Ellis, C. (2020). Semantic verbal fluency task: The effects of age, educational level, and sex in Lebanese-speaking adults. *Applied Neuropsychology: Adult* [Online ahead of print]. https://doi.org/10.1080/23279095.2020.1821031

Jebahi, F., Abou Jaoude, R., Daaboul, H., El Achkar, R., & Jacobs, M. M. (2021). Preliminary normative data for 12 categories using semantic verbal fluency: The role of animacy. *Applied*

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Neuropsychology: Adult, 1–6 [Online ahead of print]. https://doi.org/10.1080/23279095.2021.1971981

Kavé, G., & Sapir-Yogev, S. (2020). Associations between memory and verbal fluency tasks. *Journal of Communication Disorders*, *83*, 105968. https://doi.org/10.1016/j.jcomdis.2019.105968

Kim, J., Kang, Y., & Yoon, J. H. (2015). Category norms for Korean adults age 55 to 74. *Communication Sciences & Disorders*, *20*(4), 559–569. https://doi.org/10.12963/csd.15267

Kircher, T., Nagels, A., Kirner-Veselinovic, A., & Krach, S. (2011). Neural correlates of rhyming vs. lexical and semantic fluency. *Brain Research*, *1391*, 71–80. https://doi.org/10.1016/j.brainres.2011.03.054

Klumpp, H., & Deldin, P. (2010). Review of brain functioning in depression for semantic processing and verbal fluency. *International Journal of Psychophysiology*, *75*(2), 77–85. https://doi.org/10.1016/j.ijpsycho.2009.10.003

Knecht, S., Deppe, M., Dräger, B., Bobe, L., Lohmann, H., Ringelstein, E.-B., & Henningsen, H. (2000). Language lateralization in healthy right-handers. *Brain*, *123*(1), 74–81. https://doi.org/10.1093/brain/123.1.74

Kronrod, Y., Coppess, E., & Feldman, N. H. (2016). A unified account of categorical effects in phonetic perception. *Psychonomic Bulletin & Review*, *23*(6), 1681–1712. https://doi.org/10.3758/s13423-016-1049-y

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Lago, S., Scharinger, M., Kronrod, Y., & Idsardi, W. J. (2015). Categorical effects in fricative perception are reflected in cortical source information. *Brain and Language, 143*, 52–58. https://doi.org/10.1016/j.bandl.2015.02.003

Lam, L. C. W., Ho, P., Lui, V. W. C., & Tam, C. W. C. (2006). Reduced semantic fluency as an additional screening tool for subjects with questionable dementia. *Dementia and Geriatric Cognitive Disorders, 22*(2), 159–164. https://doi.org/10.1159/000094543

Ledoux, K., Vannorsdall, T. D., Pickett, E. J., Bosley, L. V., Gordon, B., & Schretlen, D. J. (2014). Capturing additional information about the organization of entries in the lexicon from verbal fluency productions. *Journal of Clinical and Experimental Neuropsychology, 3*6(2), 205–220. https://doi.org/10.1080/13803395.2013.878689

Lehtinen, N., Luotonen, I., & Kautto, A. (2021). Systematic administration and analysis of verbal fluency tasks: Preliminary evidence for reliable exploration of processes underlying task performance. *Applied Neuropsychology: Adult*, 1–13 [Online ahead of print]. https://doi.org/10.1080/23279095.2021.1973471

Linz, N., Troger, J., Alexandersson, J., Wolters, M., Konig, A., & Robert, P. (2017). Predicting dementia screening and staging scores from semantic verbal fluency performance. In R. Gottumukkala, X. Ning, G. Dong, V. Raghavan, S. Aluru, G. Karypis, … Xindong Wu (Eds.), *2017 IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 719–728). Los Alamitos, CA: Conference Publishing Services – IEEE Computer Society. https://doi.org/10.1109/icdmw.2017.100

Liouta, E., Gatzonis, S., Kalamatianos, T., Kalyvas, A., Koutsarnakis, C., Liakos, F., … Stranjalis, G. (2017). Finger tapping and verbal fluency post-tap test improvement in INPH: Its

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

value in differential diagnosis and shunt-treatment outcomes prognosis. *Acta Neurochirurgica*, *159*(12), 2301–2307. https://doi.org/10.1007/s00701-017-3301-2

Liuzzi, A. G., Aglinskas, A., & Fairhall, S. L. (2020). General and feature-based semantic representations in the semantic network. *Scientific Reports*, *10*(1), 8931. https://doi.org/10.1038/s41598-020-65906-0

Ljubešić, N., & Erjavec, T. (2011). hrWaC and slWaC: Compiling web corpora for Croatian and Slovene. In I. Habernal, & V. Matoušek (Eds.), *Text, Speech and Dialogue – 14th International Conference, TSD 2011, Pilsen, Czech Republic, September 1–5, 2011: Proceedings* (Lecture Notes in Computer Science) (Vol. 6836, pp. 395–402). Berlin: Springer. https://doi.org/10.1007/978-3-642-23538-2_50

McDonnell, M., Dill, L., Panos, S., Amano, S., Brown, W., Giurgius, S., … Miller, K. (2019). Verbal fluency as a screening tool for mild cognitive impairment. *International Psychogeriatrics*, *32*(9), 1055–1062. https://doi.org/10.1017/s1041610219000644

Mengisidou, M. (2019). *Using Verbal Fluency Tasks to Investigate the Lexicon in Greek-Speaking Children with Literacy and Language Disorders*. Ph.D. thesis: University College London, London, UK. https://discovery.ucl.ac.uk/id/eprint/10077333

Metternich, B., Buschmann, F., Wagner, K., Schulze-Bonhage, A., & Kriston, L. (2014). Verbal fluency in focal epilepsy: A systematic review and meta-analysis. *Neuropsychology Review*, *24*(2) 200–218. https://doi.org/10.1007/s11065-014-9255-8

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Meyer, G. F., Spray, A., Fairlie, J. E., & Uomini, N. T. (2014). Inferring common cognitive mechanisms from brain blood-flow lateralization data: A new methodology for fTCD analysis. *Frontiers in Psychology*, *5*, 552. https://doi.org/10.3389/fpsyg.2014.00552

Miller, E. (1984). Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. *British Journal of Clinical Psychology*, *23*(1), 53–57. https://doi.org/10.1111/j.2044-8260.1984.tb00626.x

Moreno-Martínez, F. J., Ruiz, M., & Montoro, P. R. (2017). Why almost always animals? Ranking fluency tasks for the detection of dementia based on receiver operating characteristic (ROC) and quality ROC analyses. *Dementia and Geriatric Cognitive Disorders*, *43*(1–2), 59–70. https://doi.org/10.1159/000454916

Nagels, A., Kirner-Veselinovic, A., Wiese, R., Paulus, F. M., Kircher, T., & Krach, S. (2011a). Effects of ketamine-induced psychopathological symptoms on continuous overt rhyme fluency. *European Archives of Psychiatry and Clinical Neuroscience*, *262*(5), 403–414. https://doi.org/10.1007/s00406-011-0281-8

Nagels, A., Kirner-Veselinovic, A., Krach, S., & Kircher, T. (2011b). Neural correlates of S-ketamine induced psychosis during overt continuous verbal fluency. *NeuroImage*, *54*(2), 1307–1314. https://doi.org/10.1016/j.neuroimage.2010.08.021

Nagels, A., Kircher, T., Dietsche, B., Backes, H., Marquetand, J., & Krug, A. (2012). Neural processing of overt word generation in healthy individuals: The effect of age and word knowledge. *NeuroImage*, *61*(4), 832–840. https://doi.org/10.1016/j.neuroimage.2012.04.019

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Neves, T. R. F., de Araújo, N. B., de Oliveira Silva, F., Ferreira, J. V. A., Nielsen, T. R., Engedal, K., …, & Deslandes, A. C. (2020). Accuracy of the semantic fluency test to separate healthy old people from patients with Alzheimer's disease in a low education population. *Jornal Brasileiro de psiquiatria*, *69*(2), 82–87. https://doi.org/10.1590/0047-2085000000270

Oberg, G., & Ramírez, M. (2006). Cross□linguistic meta□analysis of phonological fluency: Normal performance across cultures. *International Journal of Psychology*, *41*(5), 342–347. https://doi.org/10.1080/00207590500345872

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Pakhomov, S. V. S., Eberly, L. E., & Knopman, D. S. (2018). Recurrent perseverations on semantic verbal fluency tasks as an early marker of cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, *40*(8), 832–840. https://doi.org/10.1080/13803395.2018.1438372

Pekkala, S. (2012). Verbal fluency tasks and the neuropsychology of language. In M. Faust (Ed.), *The Handbook of the Neuropsychology of Language* (pp. 619–634). Chichester, UK: Wiley-Blackwell. https://doi.org/10.1002/9781118432501.ch30

Pendleton, M. G., Heaton, R. K., Lehman, R. A. W., & Hulihan, D. (1982). Diagnostic utility of the Thurstone Word Fluency Test in neuropsychological evaluations. *Journal of Clinical Neuropsychology*, *4*(4), 307–317. https://doi.org/10.1080/01688638208401139

Radanovic, M., Carthery-Goulart, M. T., Charchat-Fichman, H., Herrera Jr., E., Lima, E. E. P., Smid, J., Porto, C. S., & Nitrini, R. (2007). Analysis of brief language tests in the detection of

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

cognitive decline and dementia. *Dementia & Neuropsychologia*, *1*(1), 37–45. https://doi.org/10.1590/s1980-57642008dn10100007

Radanovic, M., Diniz, B. S., Mirandez, R. M., Novaretti, T. M. da S., Flacks, M. K., Yassuda, M. S., & Forlenza, O. V. (2009). Verbal fluency in the detection of mild cognitive impairment and Alzheimer's disease among Brazilian Portuguese speakers: The influence of education. *International Psychogeriatrics*, *21*(6), 1081–1087. https://doi.org/10.1017/s1041610209990639

Raucher-Chéné, D., Achim, A. M., Kaladjian, A., & Besche-Richard, C. (2017). Verbal fluency in bipolar disorders: A systematic review and meta-analysis. *Journal of Affective Disorders*, *207*, 359–366. https://doi.org/10.1016/j.jad.2016.09.039

Rosenkranz, A. (2020). Die Rolle der kognitiven Kontrolle bei der Wortverarbeitung: Hinweise für die Aphasietherapie. *Aphasie und verwandte Gebiete*, *2020*(2), 19–29. https://aphasie.org/wp-content/uploads/2021/09/APHA_Fachzeitschrift_2-2020.pdf#page=19

Rosenkranz, A., Kircher, T., & Nagels, A. (2019). Neuropsychological correlates underlying verbal fluency deficits in schizophrenia: The role of attention and executive function. In V. Erdeljac, & M. Sekulić Sović (Eds.), *Interdisciplinary Linguistic and Psychiatric Research on Language Disorders* (pp. 45–54). Zagreb: FF Press. https://doi.org/10.17234/9789531758314.04

Santangelo, G., Altieri, M., Enzinger, C., Gallo, A., & Trojano, L. (2019). Cognitive reserve and neuropsychological performance in multiple sclerosis: A meta-analysis. *Neuropsychology*, *33*(3), 379–390. https://doi.org/10.1037/neu0000520

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Scheuringer, A., Harris, T.-A., & Pletzer, B. (2020). Recruiting the right hemisphere: Sex differences in inter-hemispheric communication during semantic verbal fluency. *Brain and Language*, *207*, 104814. https://doi.org/10.1016/j.bandl.2020.104814

Seijdel, N., Scholte, H. S., & de Haan, E. H. F. (2021). Visual features drive the category-specific impairments on categorization tasks in a patient with object agnosia. *Neuropsychologia*, *161*, 108017. https://doi.org/10.1016/j.neuropsychologia.2021.108017

Sekulić Sović, M., Erdeljac, V., Kužina, I., Vandek, M., Mimica, N., Ostojić, D., & Savić, A. (2019). Shared lexical-semantic features and the animacy effect in early-course psychosis. In V. Erdeljac, & M. Sekulić Sović (Eds.), *Interdisciplinary Linguistic and Psychiatric Research on Language Disorders* (pp. 69–81). Zagreb: FF Press. https://doi.org/10.17234/9789531758314.06

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*, 772. https://doi.org/10.3389/fpsyg.2014.00772

Spek, A., Schatorjé, T., Scholte, E., & van Berckelaer-Onnes, I. (2009). Verbal fluency in adults with high functioning autism or Asperger syndrome. *Neuropsychologia*, *47*(3), 652–656. https://doi.org/10.1016/j.neuropsychologia.2008.11.015

Stielow, A. (2017). Semantische und formal-lexikalische Wortgenerierungsleistungen bei Kindern, Jugendlichen und jungen Erwachsenen. *Forschung Sprache*, *5*(2), 129–139. https://www.forschung-sprache.eu/fileadmin/user_upload/Dateien/Heftausgaben/2017-2/FS_2_2017_Stielow_129.pdf

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Stielow, A., & Stenneken, P. (2017). Der Beitrag exekutiver und sprachlicher Teilleistungen in semantischen und formal-lexikalischen Wortflüssigkeitsaufgaben. *Sprache · Stimme · Gehör, 41*(04), 197–203. https://doi.org/10.1055/s-0043-117202

Sutin, A. R., Stephan, Y., & Terracciano, A. (2019). Verbal fluency and risk of dementia. *International Journal of Geriatric Psychiatry, 34*(6), 863–867. https://doi.org/10.1002/gps.5081

Tagini, S., Seyed-Allaei, S., Scarpina, F., Toraldo, A., Mauro, A., Cherubini, P., & Reverberi, C. (2021). When fruits lose to animals: Disorganized search of semantic memory in Parkinson's disease. *Neuropsychology, 35*(5), 529–539. https://doi.org/10.1037/neu0000429

Takizawa, R., Fukuda, M., Kawasaki, S., Kasai, K., Mimura, M., Pu, S., … Okazaki, Y. (2014). Neuroimaging-aided differential diagnosis of the depressive state. *NeuroImage, 85*(Part 1), 498–507. https://doi.org/10.1016/j.neuroimage.2013.05.126

Tan, E. J., Neill, E., Tomlinson, K., & Rossell, S. L. (2020). Semantic memory impairment across the schizophrenia continuum: A meta-analysis of category fluency performance. *Schizophrenia Bulletin Open, 1*(1), sgaa054. https://doi.org/10.1093/schizbullopen/sgaa054

Tan, E. J., Neill, E., Tomlinson, K., & Rossell, S. L. (2021). Corrigendum to: Semantic memory impairment across the schizophrenia continuum: A meta-analysis of category fluency performance. *Schizophrenia Bulletin Open, 2*(1), sgab018. https://doi.org/10.1093/schizbullopen/sgab018

Tassi, E., Boscutti, A., Mandolini, G. M., Moltrasio, C., Delvecchio, G., & Brambilla, P. (2022). A scoping review of near infrared spectroscopy studies employing a verbal fluency task in

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

bipolar disorder. *Journal of Affective Disorders*, *298*(Part A), 604–617. https://doi.org/10.1016/j.jad.2021.11.019

Thiele, K., Quinting, J. M., & Stenneken, P. (2016). New ways to analyze word generation performance in brain injury: A systematic review and meta-analysis of additional performance measures. *Journal of Clinical and Experimental Neuropsychology*, *38*(7), 764–781. https://doi.org/10.1080/13803395.2016.1163327

Tröger, J., Linz, N., König, A., Robert, P., & Alexandersson, J. (2018). Telephone-based dementia screening I: Automateds semantic verbal fluency assessment. In: *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare* (pp. 59–66). Association for Computing Machinery. https://doi.org/10.1145/3240925.3240943

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. https://doi.org/10.1037//0894-4105.11.1.138

Troyer, A. K., & Moscovitch, M. (2006). Cognitive processes of verbal fluency tasks. In A. M. Poreh (Ed.), *The Quantified Process Approach to Neuropsychological Assessment* (pp. 143–160). New York, NY: Taylor & Francis. https://www.taylorfrancis.com/chapters/mono/10.4324/9780203720899-20

Turkstra, L. S. (2018). Verbal fluency. In J. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 3576–3579). Cham: Springer. https://doi.org/10.1007/978-3-319-56782-2_1423-3

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Turkstra, L. S., Coelho, C., & Ylvisaker, M. (2005). The use of standardized tests for individuals with cognitive-communication disorders. *Seminars in Speech and Language*, *26*(04), 215–222. https://doi.org/10.1055/s-2005-922101

Uomini, N. T., & Meyer, G. F. (2013). Shared brain lateralization patterns in language and Acheulean stone tool production: A functional transcranial Doppler ultrasound study. *PLoS ONE*, *8*(8), e72693. https://doi.org/10.1371/journal.pone.0072693

Vandek, M., Gabrić, P., Kužina, I., Erdeljac, V., & Sekulić Sović, M. (2018, July 2–4). Verbal fluency and working memory interaction [Conference poster]. *10th International Workshop on Language Production*, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands. https://doi.org/10.5281/zenodo.4681276

Vaughan, R. M., Coen, R. F., Kenny, R., & Lawlor, B. A. (2018). Semantic and phonemic verbal fluency discrepancy in mild cognitive impairment: Potential predictor of progression to Alzheimer's disease. *Journal of the American Geriatrics Society*, *66*(4), 755–759. https://doi.org/10.1111/jgs.15294

Ventura, P., Morais, J., & Kolinsky, R. (2005). Evaluating feature-category relations using semantic fluency tasks. *Brain and Cognition*, *58*(2), 202–212. https://doi.org/10.1016/j.bandc.2004.11.004

Vicente, S. G., Benito-Sánchez, I., Barbosa, F., Gaspar, N., Dores, A. R., Rivera, D., & Arango-Lasprilla, J. C. (2021). Normative data for Verbal Fluency and Object Naming Tests in a sample of European Portuguese adult population. *Applied Neuropsychology: Adult*, 1–12 [Online ahead of print]. https://doi.org/10.1080/23279095.2020.1868472

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Vonk, J. M. J., Rizvi, B., Lao, P. J., Budge, M., Manly, J. J., Mayeux, R., & Brickman, A. M. (2019). Letter and category fluency performance correlates with distinct patterns of cortical thickness in older adults. *Cerebral Cortex*, *29*(6), 2694–2700. https://doi.org/10.1093/cercor/bhy138

Wende, K. C., Straube, B., Stratmann, M., Sommer, J., Kircher, T., & Nagels, A. (2012). Neural correlates of continuous causal word generation. *NeuroImage*, *62*(3), 1399–1407. https://doi.org/10.1016/j.neuroimage.2012.06.003

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult*, *23*(1), 29–34. https://doi.org/10.1080/23279095.2015.1004574

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias (Eds.), *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)* (pp. 1556–1559). Genoa: European Language Resources Association. http://www.lrec-conf.org/proceedings/lrec2006/pdf/153_pdf.pdf

Wong, M. N., Murdoch, B., & Whelan, B.-M. (2010). Language disorders subsequent to mild traumatic brain injury (MTBI): Evidence from four cases. *Aphasiology*, *24*(10), 1155–1169. https://doi.org/10.1080/02687030903168212

Yeung, M. K., & Lin, J. (2021). Probing depression, schizophrenia, and other psychiatric disorders using fNIRS and the verbal fluency test: A systematic review and meta-analysis. *Journal of Psychiatric Research*, *140*, 416–435. https://doi.org/10.1016/j.jpsychires.2021.06.015

medRxiv preprint doi: https://doi.org/10.1101/2021.12.30.21268567; this version posted January 26, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Zhao, Q., Guo, Q., & Hong, Z. (2013). Clustering and switching during a semantic verbal fluency test contribute to differential diagnosis of cognitive impairment. *Neuroscience Bulletin*, *29*(1), 75–82. https://doi.org/10.1007/s12264-013-1301-7