



Graph analysis of verbal fluency test discriminate between patients with Alzheimer's disease, mild cognitive impairment and normal elderly controls

Laiss Bertola^{1*†}, Natália B. Mota^{2†}, Mauro Copelli³, Thiago Rivero¹, Breno Satler Diniz^{1,4}, Marco A. Romano-Silva^{4,5}, Sidarta Ribeiro² and Leandro F. Malloy-Diniz^{1,4}

¹ Laboratory of Clinical Neuroscience Investigations, Federal University of Minas Gerais, Belo Horizonte, Brazil

² Brain Institute, Federal University of Rio Grande do Norte, Natal, Brazil

³ Physics Department, Federal University of Pernambuco, Recife, Brazil

⁴ Mental Health Department, Faculty of Medicine, Federal University of Minas Gerais, Belo Horizonte, Brazil

⁵ Faculty of Medicine, National Institute of Science and Technology – Molecular Medicine, Federal University of Minas Gerais, Belo Horizonte, Brazil

Edited by:

Manuel Menéndez-González,
Hospital Álvarez-Buylla, Spain

Reviewed by:

Roberta Brinton, University of
Southern California, USA

Douglas Watt, Quincy Medical
Center, USA

Mikhail Lebedev, Duke University,
USA

*Correspondence:

Laiss Bertola, Laboratory of Clinical
Neuroscience Investigations,
Faculty of Medicine, Federal
University of Minas Gerais, Av.
Alfredo Balena, 190 office 235, Belo
Horizonte, Minas Gerais,
CEP 30.130-100, Brazil
e-mail: laissbertola@gmail.com

[†] These authors contributed equally
to this study.

Verbal fluency is the ability to produce a satisfying sequence of spoken words during a given time interval. The core of verbal fluency lies in the capacity to manage the executive aspects of language. The standard scores of the semantic verbal fluency test are broadly used in the neuropsychological assessment of the elderly, and different analytical methods are likely to extract even more information from the data generated in this test. Graph theory, a mathematical approach to analyze relations between items, represents a promising tool to understand a variety of neuropsychological states. This study reports a graph analysis of data generated by the semantic verbal fluency test by cognitively healthy elderly (NC), patients with Mild Cognitive Impairment—subtypes amnesic (aMCI) and amnesic multiple domain (a+mdMCI)—and patients with Alzheimer's disease (AD). Sequences of words were represented as a speech graph in which every word corresponded to a node and temporal links between words were represented by directed edges. To characterize the structure of the data we calculated 13 speech graph attributes (SGA). The individuals were compared when divided in three (NC—MCI—AD) and four (NC—aMCI—a+mdMCI—AD) groups. When the three groups were compared, significant differences were found in the standard measure of correct words produced, and three SGA: diameter, average shortest path, and network density. SGA sorted the elderly groups with good specificity and sensitivity. When the four groups were compared, the groups differed significantly in network density, except between the two MCI subtypes and NC and aMCI. The diameter of the network and the average shortest path were significantly different between the NC and AD, and between aMCI and AD. SGA sorted the elderly in their groups with good specificity and sensitivity, performing better than the standard score of the task. These findings provide support for a new methodological frame to assess the strength of semantic memory through the verbal fluency task, with potential to amplify the predictive power of this test. Graph analysis is likely to become clinically relevant in neurology and psychiatry, and may be particularly useful for the differential diagnosis of the elderly.

Keywords: semantic verbal fluency, graph analysis, elderly, Alzheimer's disease, mild cognitive impairment

INTRODUCTION

Language and semantic memory tend to remain stable across the human lifespan in contrast to other cognitive domains, like episodic memory and attention, which usually decline after the 5th decade (Craik and Bialystok, 2006). They are also usually spared in the initial stages of neurodegenerative disorders, such as Alzheimer's disease (AD), though we can still observe milder deficits, e.g., anomia or reduced semantic verbal fluency, which can be identified in a comprehensive neuropsychological evaluation (Henry et al., 2004; Garrard

et al., 2005; Nutter-Upham et al., 2008; Taler and Phillips, 2008).

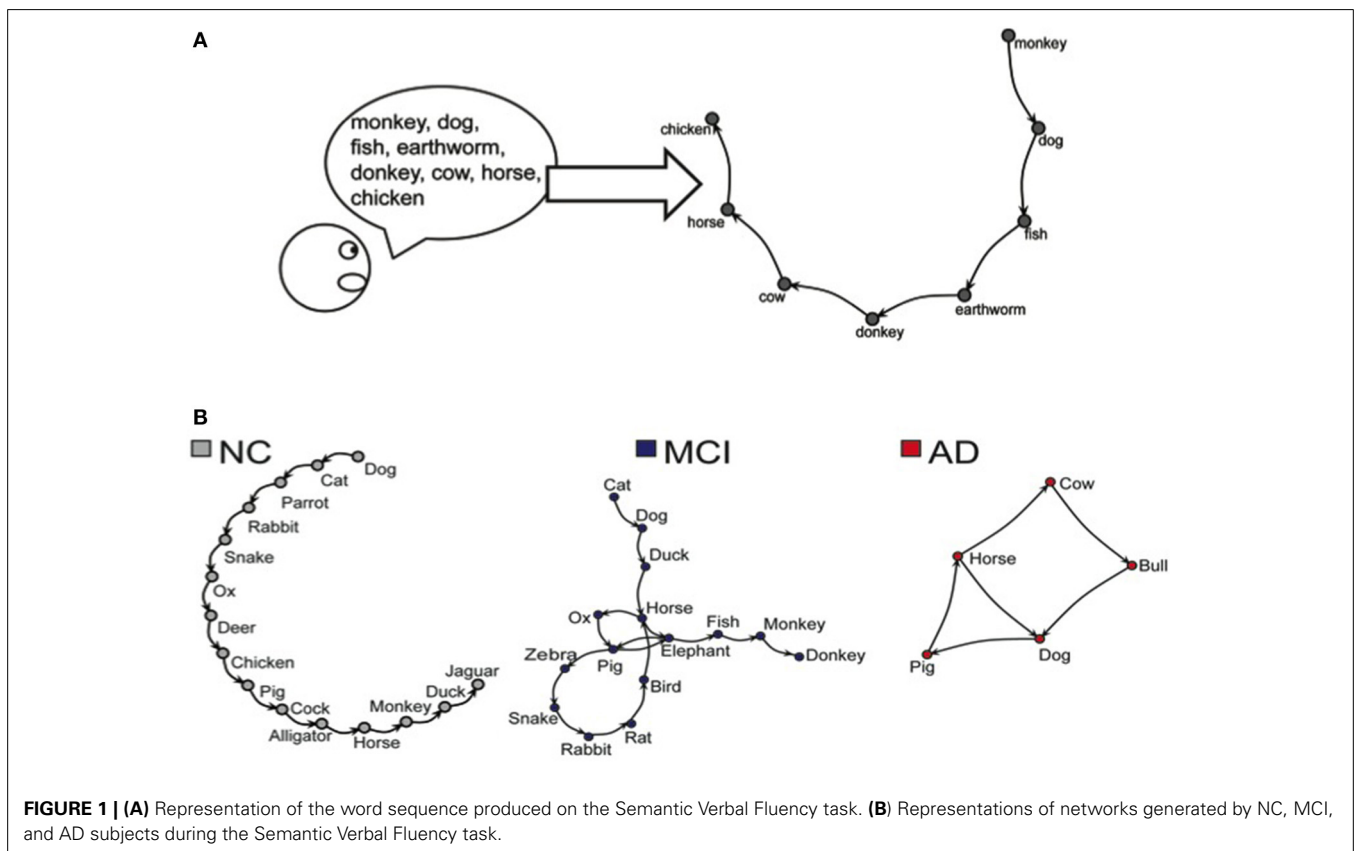
Verbal fluency is the ability to produce a satisfying sequence of spoken words during a given time interval (Nickles, 2001). Verbal fluency tests are experimentally designed to assess this ability through the production of words starting with a specific letter (Phonemic Verbal Fluency) or belonging to a category of knowledge (Semantic Verbal Fluency). Semantic verbal fluency is one of the most commonly used tasks to evaluate language and semantic memory skills in older adults. This task depends on

the preservation of language (e.g., words can be spoken correctly during the task), though it is significantly influenced by semantic memory (e.g., the knowledge of the category asked must be intact) and executive function (e.g., the ability to search the asked knowledge) domains (Adlam et al., 2006; Unsworth et al., 2011). This task often activates the temporal lobe, a region broadly related to conceptualization, general information and knowledge about names (Patterson et al., 2007). Semantic verbal fluency contributes to predict future cognitive and functional impairments in the elderly (Salmon et al., 2002; Amieva et al., 2005; Hodges et al., 2006; Aretouli et al., 2011), and predict the progression from MCI to AD (Saxton et al., 2004).

Despite being widely used for neuropsychological assessment in the elderly, the standard measure of the verbal fluency test is restricted to the total of correct words produced in the task (Lezak et al., 2004; Strauss et al., 2006), and does not take into account other clinically-relevant information that may be contained in the patient's specific performance. This task requires the production of words belonging to a specific category, and each subject produced the words following an order of exemplars during the 1-min task. This order of words produced allows the construction of a network based on the temporal link between the words. These temporal links may inform that words produced in a specific temporal sequence are probably conceptually related, as suggested by the semantic association models (McClelland and Rogers, 2003; Griffiths et al., 2007).

Goni et al. (2010) constructed a semantic network using the verbal fluency task applied to an adult sample, and represented the semantic memory as a graph ruled by conceptual constraints. A normal semantic verbal fluency network is represented by a directed graph with only one occurrence for each word. Lerner et al. (2009) investigated the network properties of subjects with MCI and AD, and found that the path lengths of the network decline while the clustering coefficient increases in the MCI and AD subjects compared to healthy elderly controls. These results showed that the normal characteristics of the semantic verbal network are significantly changed in the continuum from normal aging to AD.

The analysis of network properties helps understanding the dynamics and organization of the cognitive and behavioral processes. A graph represents a network with nodes linked by edges (Mota et al., 2012). Formally, a graph is a mathematical representation of a network $G = (N, E)$, with $N = \{w_1, w_2, \dots, w_n\}$ a set of nodes and $E = \{(w_i, w_j)\}$ a set of edges or links between words w_i in N and w_j in N . The interpretation of the meaning of a graph depends on what is being represented (Butts, 2009; Mota et al., 2012). We carried out an analysis of the network properties of the semantic verbal fluency of subjects with MCI or AD. We hypothesize that the analysis of the semantic verbal fluency network properties can help to better discriminate between older adults with normal cognitive performance, mild cognitive impairment or Alzheimer's disease. This approach had been used



with success to identify patients with schizophrenia and bipolar disorder (Mota et al., 2012, 2014).

MATERIALS AND METHODS

SUBJECTS

One hundred older adults were included in this study. All subjects were assessed in the Centro de Referência à Saúde do Idoso Jenny de Andrade Faria, Clinical Hospital, Federal University of Minas Gerais. All the participants underwent a comprehensive clinical and neuropsychological assessment. The neuropsychological protocol included the following tests: Mini Mental State Exam, Frontal Assessment Battery, Category Verbal Fluency of Animals and Fruits, Letter Fluency of S, Digit Span, Stick Design Test, Clock Drawing Test, Rey Auditory Verbal Learning Test, Naming Test (TN-LIN), and Token Test. This protocol has been validated for the neuropsychological assessment of older adults with low educational status (de Paula et al., 2013). After the clinical and neuropsychological assessment, and adjudication meeting was held and the final diagnosis was reached by consensus. The AD diagnosis was based on the proposed criteria of McKhann et al. (1984) and the patient should present general and worsening cognitive impairment, in two or more cognitive domains, and functional impairment in the daily living activities. The MCI diagnosis followed the criteria proposed by Winblad et al. (2004), where the older adult presents cognitive decline in one or more cognitive domains but is preserved in basic and instrumental daily living activities or presents a minimal impairment. The MCI subgroup division considered the amnesic MCI (aMCI) classification for participants that only present memory impairment, and amnesic multiple-domain MCI (a+mdMCI) for participants that present impairment in memory and other cognitive domain, though fulfilling all the MCI criteria established by Winblad et al. (2004).

The project was approved by the Research Ethics Committee of the Federal University of Minas Gerais (COEP-334/06). The subjects were divided into four groups: (1) normal cognitive performance (NC), $n = 25$; (2) amnesic single-domain MCI (aMCI), $n = 25$; (3) amnesic multiple-domain MCI (a+mdMCI), $n = 25$; (4) AD, $n = 25$.

VERBAL FLUENCY TEST

The participants performed the Semantic Verbal Fluency test, category of animals, for which they were asked to produce the maximum names of animals within 60 s; explicit/implicit instructions were given to avoid repetitions. All the words were recorded, including repetitions and errors. The scoring procedure included: total of words produced, total of correct words, total of errors, total of repetitions, and the fraction of repetitions according to the total of words produced by each participant. The scores in this task were not taken into account in the diagnosis adjudication of each participant.

STATISTICAL ANALYSIS

The study design involved two stages of analysis, considering three (NC, MCI, AD) or four groups (NC, aMCI, a+mdMCI, AD), and the same statistical analysis and graph measures were performed for comparing the three or four groups. The

MCI group comprised both the aMCI and the a+mdMCI groups.

We performed the Shapiro-Wilk test of normality of the sample, and since the majority of the variables did not fit the assumption of normality, we used the Kruskal-Wallis test of differences between several independent groups and the Wilcoxon Rank sum test for two independent samples. Bonferroni correction was applied to all analyses.

Group sorting was implemented with a Naïve Bayes classifier, which shows superior performance with small samples (Singh and Provan, 1995; Kotsiantis, 2007). The choice of attributes for the classifier was based on significant correlations of the attributes with established clinical measures of differential diagnosis (global cognitive status and daily living functionality). Sensitivity, specificity and the area under the receiver operating characteristic curve (AUC) were used to estimate classification quality, which was considered excellent when AUC was higher than 0.8, good when AUC ranged from 0.6 to 0.8, and poor (not above the chance), when AUC was smaller than 0.6.

GRAPH MEASURES

The word sequence produced on the Semantic Verbal Fluency test was represented as a speech graph, using the software *SpeechGraphs* (Mota et al., 2014). The program represents a text (in this case, the sequence of words produced by the verbal fluency test) as a graph, representing every word as a node, and the temporal link between words as an edge (Figure 1).

We then calculated word count (WC) and 13 additional Speech Graph Attributes (SGA) comprising general attributes: total of nodes (N) and edges (E); connected components: the largest strongly connected component (LSC); recurrence attributes: repeated (RE) and parallel edges (PE), cycles of one (L1), two (L2), or 3 nodes (L3); global attributes: average total degree (ATD), density, diameter, average shortest path (ASP) and clustering coefficient (CC) (for more detailed information see Supplementary Table and Figure on Supplementary Material).

Given the task instructions, we expected the subjects to produce a linear network, i.e., a sequence in which each correct word

Table 1 | Socio-demographic data, verbal fluency and Speech Graph Attributes of NC, MCI, and AD groups, with Bonferroni-corrected significant differences across groups established by the Kruskal-Wallis comparison.

	NC			MCI			AD			<i>p</i>
	Median	IQR		Median	IQR		Median	IQR		
		Q1	Q3		Q1	Q3		Q1	Q3	
Age	76	72	80	76	71	81	78	67	81	0.9785
Education	4	3	4	4	2	4	4	3	4	0.8400
Katz	0	0	0	–	–	–	0	0	0	0.0105
Lawton	0	0	0	0	0	1	6	4	8	0.0000
MMSE	27	24	29	25	23	27	20	17	23	0.0000

Katz, Katz Index; Lawton, Lawton Index; MMSE, Mini Mental State Exam; IQR, Interquartile Range; Q1, 1th Quartile; Q3, 3rd Quartile. Red values have significance $p = 0.0167$.

was followed by a different correct word, without repetitions. A correct performance in this test should yield graphs with identical number of nodes (N) and words (WC), N-1 edges, no recurrence (i.e., without parallel edges, repeated edges or loops), and zero strongly connected components (LSC). In addition, the average total degree (ATD) should be close to 2, with a very small density,

very low clustering coefficient (CC), and large distances (diameter should be equal to E).

RESULTS

Table 1 shows data for socio-demographic data, Mini Mental State Exam (MMSE), total number of produced words in the

Table 2 | Verbal fluency and Speech Graph Attributes of NC, MCI, and AD groups, with Bonferroni-corrected significant differences across groups established by the Kruskal-Wallis comparison.

	NC			MCI			AD			p
	Median	IQR		Median	IQR		Median	IQR		
		Q1	Q3		Q1	Q3		Q1	Q3	
VFE	0	0	0	0	0	0	0	0	0	1.0000
VFPR	0	0	0.07	0	0	0.13	0	0	0.1	0.2330
VFR	0	0	1	0	0	1	0	0	1	0.4462
VFC	14	12	15	11	10	14	9	7	10	0.0000
VFTT	15	13	15	12	10	15	9	8	10	0.0000
WC	15	13	15	12	10	15	9	8	10	0.0000
N	14	13	15	11	10	14	9	7	10	0.0000
E	14	12	14	11	9	14	8	7	9	0.0000
RE	–	–	–	0	0	0	0	0	0	0.6034
PE	0	0	0	0	0	0	0	0	0	0.6591
L1	–	–	–	0	0	0	–	–	–	0.6065
L2	0	0	0	0	0	0	0	0	0	0.6942
L3	0	0	0	0	0	0	0	0	0	0.0265
LSC	1	1	7	1	1	6	1	1	4	0.7568
ATD	1.86	1.85	2.00	1.87	1.82	2.00	1.80	1.75	2.00	0.2584
Diameter	12	9.00	13.00	9	6.00	12.00	7	5.00	8.00	0.0001
ASP	4.66	3.67	5.20	3.66	2.91	4.67	3	2.29	3.33	0.0001
CC	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00	0.2479
Density	0.07	0.06	0.08	0.08	0.07	0.10	0.10	0.10	0.14	0.0000

VFE, errors; VFPR, percentage of repetitions; VFR, repetitions; VFC, corrects words; VFTT, total of words; WC, word count; N, nodes; E, edges; RE, repeated; PE, parallel edges; L1, L2, L3, cycles of one, two or 3 nodes; LSC, largest strongly connected component; ATD, average total degree; ASP, average shortest path; CC, clustering coefficient. Red values have significance $p = 0.0167$.

Table 3 | Pairwise group comparison with Bonferroni-corrected significant differences between groups established by Wilcoxon Ranksum test.

	NC x MCI			NC x AD			MCI x AD		
	W	Z	p	W	z	p	W	Z	p
Katz	1875.0	1.385	0.1658	598.5	–1.388	0.1651	1800.0	2.815	0.0049
Lawton	707.0	–3.396	0.0007	325.0	–6.254	0.0000	1327.5	6.358	0.0000
MMSE	1712.5	2.114	0.0345	394.0	4.732	0.0000	558.5	–4.412	0.0000
VFC	1626.0	3.088	0.0020	396.5	4.776	0.0000	658.5	–3.427	0.0006
VFTT	1689.5	2.375	0.0175	410.5	4.513	0.0000	656.5	–3.432	0.0006
WC	1689.5	2.375	0.0175	410.5	4.513	0.0000	656.5	–3.432	0.0006
N	1626.0	3.091	0.0020	394.0	4.824	0.0000	640.5	–3.629	0.0003
E	1689.5	2.375	0.0175	410.5	4.513	0.0000	656.5	–3.432	0.0006
L3	1862.5	1.225	0.2205	600.0	–1.137	0.2552	1787.5	2.613	0.0090
Diameter	1686.5	2.405	0.0161	425.5	4.238	0.0000	720.0	–2.783	0.0054
ASP	1667.0	2.618	0.0088	414.5	4.433	0.0000	720.0	–2.773	0.0055
Density	653.0	–3.338	0.0008	388.5	–4.924	0.0000	1596.5	3.558	0.0004

W, Wilcoxon Ranksum; z, = z score. Red values have significance $p = 0.0167$.

verbal fluency test, total number of correct words produced, total number of repetitions performed during the task, the percentage of repetitions performed according to the total of produced words, and the errors produced.

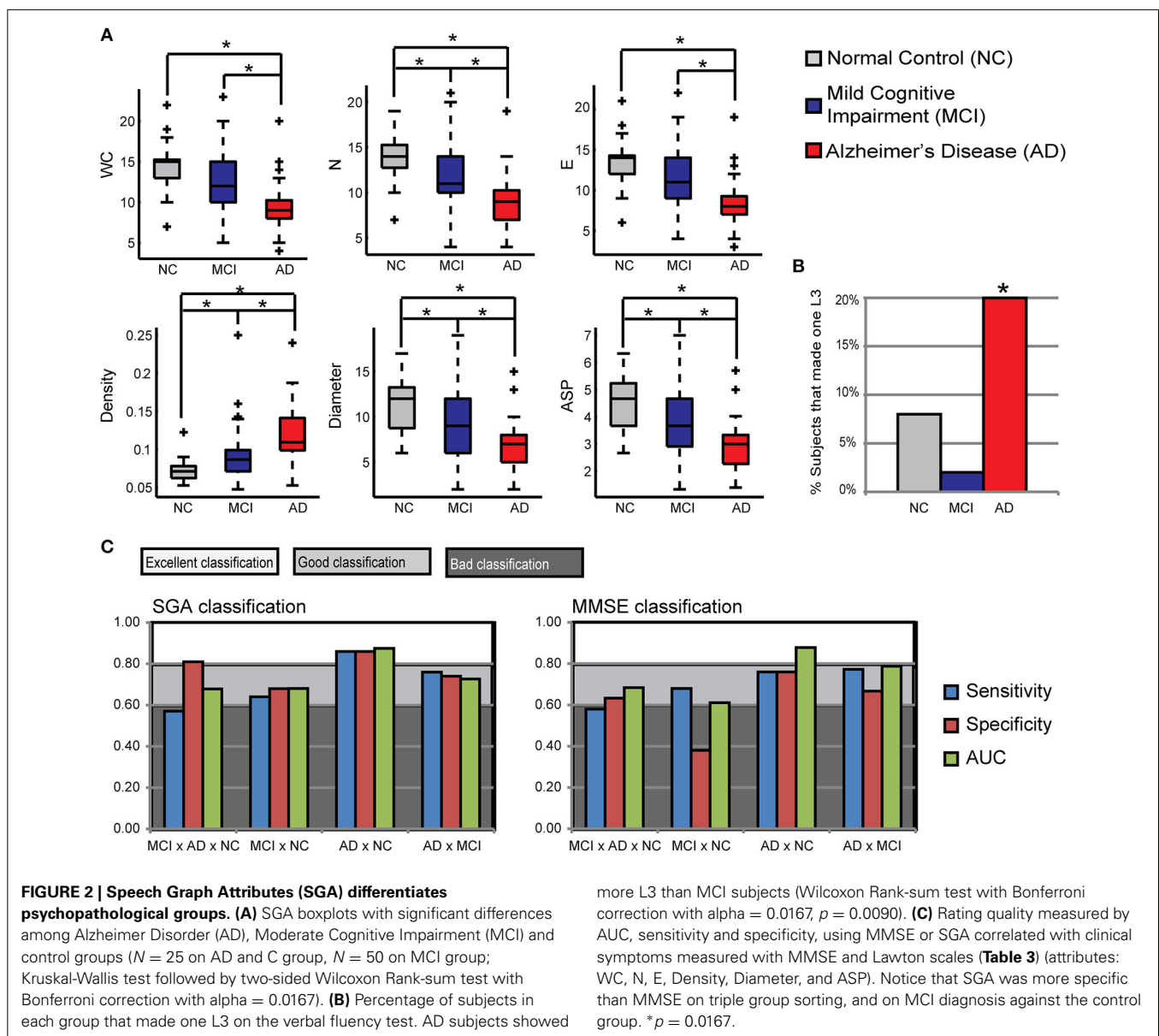
The groups did not differ in age and education, and only the control group had a significant difference in gender distributions ($X^2 = 6.76$, $df = 2$, $p = 0.009$) (Table 1). The results of the groups' comparison on the daily living activities, the global cognitive status are also reported on Table 1. Verbal fluency measures and the Speech Graph Attributes are reported on Table 2.

Despite the lower number of correct words produced by the NC group, it is similar to those observed to Brazilian normative data (Brucki et al., 1997). Moreover, the scores on the verbal fluency test were not taken into account for participant classification into the diagnostic groups.

The groups significantly differed in the performance on ADLs, in general cognitive status, number of correct words and total words produced, and in the Speech Graph measures of word count, nodes, edges, loops of 3 nodes, diameter, average short path and density. As expected, the NC group performed better at ADLs, had higher scores on the MMSE, produced more nodes, a network with larger diameter and less dense, when compared with the MCI and AD groups. The MCI group showed an intermediate performance between NC and AD groups in all measures.

Table 3 and Figure 2A show pairwise comparisons of the 3 diagnosis groups. Statistical significance was set at $p < 0.0167$, after Bonferroni correction for multiple comparisons.

The comparison of the variables between NC and MCI groups demonstrate that the groups differ in the index of instrumental daily living activities, in the number of correct words produced,



number of nodes, diameter, average short path and density of the network. The NC produced less dense graphs with more nodes, and larger Diameter and ASP than the MCI and AD. Furthermore, NC made more edges, total words produced, and had a better general cognitive status than the AD group. The MCI and AD groups differ in all measures, demonstrating that a change in the general cognitive status, functionality, verbal fluency measures and the speech graph attributes (WC, N, E, L3, Diameter, ASP, and Density) (**Figure 2B**) almost follow a continuous modification as the diagnosis impairs.

Table 4 shows the Spearman correlations between the SGA and the clinical measures of differential diagnosis (global cognitive status—MMSE—and daily living functionality—Katz and Lawton Index). The significance level was established in $p = 0.0012$ after a Bonferroni correction for 42 comparisons.

We found significant correlations between the MMSE and the SGA of Nodes and Density, indicating that the more cognitively preserved elderly produced a larger number of unique nodes, and networks with a smaller density than cognitively impaired subjects. The correlation between the attributes and the Lawton Index of instrumental daily living activities revealed that the more functionally dependent were the elderly, the less words, nodes and edges they produced, showing networks with a smaller diameter and average short path, but a higher density. These results indicate that functional autonomy correlate more with SGA than with the general cognitive status.

The Naïve Bayes classifier results (**Figure 2C**) show that a selection of SGA correlated with functional and cognitive impairment measured by other instruments, provided good to excellent classification power, being similar to the MMSE classification power, or even better for the distinction between the NC and MCI groups. When the SGA were associated to the Lawton Index or the MMSE, the power of classification increased; a combination of the 3 measurements provided maximal classification quality (**Table 5**). Overall, the combination of graph measures and functional

dependence yielded very accurate differential classification of the AD (1.00) and MCI (0.78) against the NC group, and between the MCI and AD (0.84).

The additional description of the socio-demographic data, Mini Mental State Exam (MMSE), verbal fluency measures of the two subgroups of MCI are reported on **Table 6**, and also the results of the four groups' comparison on the sociodemographic variables. **Table 7** shows the four group comparison on the verbal fluency and Speech Graph Attributes.

A comparison of the four groups showed significant differences in daily functionality, general cognitive status, total and correct words produced, and in the SGA word count, nodes, edges, diameter, ASP and density (same attributes found in the three-group comparison).

Table 8 and **Figure 3A** compare the four groups of elderly, with Bonferroni correction for multiple comparisons ($\alpha = 0.0083$).

The pairwise comparison detected no significant differences between MCI subtypes in the measures selected in this study. The difference between the NC and aMCI groups occurred only in instrumental daily living functionality, i.e., NC are more independent than aMCI. The significant differences between the NC and AD and between aMCI and AD are similar; the NC and aMCI groups are less functionally dependent, have better cognitive status, produce more total and correct words, a higher word count, more nodes and edges, higher Diameter and ASP, and less dense networks when compared to the AD group. The NC are more functionally independent, produce more total and correct words, a higher word count, more nodes and edges, and a network less dense than the a+mdMCI group. AD subjects, comparable to the a+mdMCI group, were more functionally dependent, showed general cognitive impairment, and produced fewer nodes and a denser network.

The Naïve Bayes classifier results (**Figure 3B**) indicate that the selected SGA has a good classification power to the diagnosis of MCI subtypes against cognitive healthy aging, and also a good classification against the dementia group. On the other hand, SGA yielded a poor classification when used to distinguish between the two subtypes of MCI. When SGA were combined with the Lawton Index, we observed an increase in the power of classification across the four groups, except between the two MCI subtypes.

The combination of the SGA with the MMSE, showed less power when compared to the combination with the Lawton index; the combination of these three variables barely improved the classification beyond the SGA and Lawton combination. These results indicate that the combination of graph measures and functional dependence again provides for good classification across the three groups ($AUC = 0.71-0.85$), except between the MCI subtypes ($AUC = 0.47$).

DISCUSSION

The aim of the present study was to assess graph-theoretical differences in the execution of a verbal fluency task among elderly with normal and pathological aging. Our results demonstrate that SGA differed significantly among the AD, MCI, and NC groups and it could be used to classify the groups. The present results

Table 4 | Spearman correlation (RHO and p -values) between SGA scores and the Katz, Lawton or MMSE scores.

	Katz		Lawton		MMSE	
	RHO	P	RHO	P	RHO	P
WC	-0.1762	0.0796	-0.4519	0.0000	0.3161	0.0014
N	-0.1811	0.0714	-0.4963	0.0000	0.3335	0.0007
E	-0.1762	0.0796	-0.4519	0.0000	0.3161	0.0014
RE	0.3014	0.0023	0.0698	0.4900	-0.0050	0.9609
PE	0.1031	0.3075	-0.0239	0.8137	0.0958	0.3432
L1	-0.0230	0.8199	-0.0888	0.3797	-0.0943	0.3505
L2	-0.0579	0.5670	-0.0673	0.5062	0.1116	0.2690
L3	0.1048	0.2993	0.2737	0.0059	-0.1557	0.1219
LSC	-0.0349	0.7305	-0.0545	0.5905	0.1171	0.2458
ATD	-0.0352	0.7279	-0.1220	0.2267	0.1611	0.1094
Diameter	-0.1339	0.1842	-0.3897	0.0001	0.2433	0.0147
ASP	-0.1480	0.1418	-0.4017	0.0000	0.2549	0.0105
CC	0.0692	0.4941	0.1786	0.0755	-0.1379	0.1713
Density	0.1766	0.0788	0.4933	0.0000	-0.3239	0.0010

Red values have significance $p = 0.0012$.

Table 5 | Rating quality measured by AUC, using SGA (attributes: WC, N, E, Density, Diameter, and ASP) correlated with clinical symptoms measured with MMSE and Lawton scales, in addition with Lawton, MMSE or both, classifying AD and MCI from NC, AD from MCI, and also classifying subtypes of MCI (aMCI or a+mdMCI) from NC or AD, or from each other.

		SGA	SGA+MMSE	SGA+ Lawton	SGA+MMSE+Lawton	MMSE	Lawton
NC x	MCI	0.681	0.716	0.780	0.793	0.638	0.649
	aMCI	0.619	0.618	0.710	0.714	0.586	0.612
	a+mdMCI	0.710	0.738	0.803	0.822	0.694	0.746
	AD	0.875	0.886	1.000	1.000	0.888	1.000
aMCI x	a+mdMCI	0.486	0.470	0.472	0.483	0.631	0.494
	AD	0.767	0.824	0.856	0.856	0.854	0.957
a+mdMCI x	AD	0.652	0.717	0.814	0.811	0.772	0.959
MCI x	AD	0.727	0.793	0.849	0.858	0.813	0.958

Table 6 | Additional description of socio-demographic data for the MCI subtypes, and the four groups comparison.

	aMCI		a+mdMCI			<i>p</i>
	Median	IQR	Median	IQR		
		Q1	Q3	Q1	Q3	
Age	75	71	79	79	81	0.7561
Education	4	2	5	3	4	0.4662
Katz	0	–	–	0	–	0.0279
Lawton	0	0	1	1	0	0.0000
MMSE	26	23	28	24	23	0.0000

*p** group comparison (NC; aMCI; a+mdMCI; AD). Red values have significance *p* = 0.0083.

show the potential of graph analysis of verbal fluency task to discriminate between these groups in clinical practice.

The correlation between the SGA and the MMSE or the Lawton Index indicate that the SGA are associated with the general cognitive status and functional performance, two important clinical measures used in geriatric assessment. Patients with worse scores in the MMSE produced fewer numbers of nodes and a less dense network. As the functional performance decreases, indicating more severe cognitive impairment stages, the networks became denser, with a smaller diameter and average short path and with fewer numbers of nodes and edges. Their networks became smaller in the number of words, with a small path through the first word to the last one, and their animals have more connection with different neighbors than would be necessary. Subjects more cognitively impaired tended to perform more dependently on their daily activities. Importantly, some attributes of SGA could indicate the progression of cognitive impairment and functional decline, as shown by denser and smaller networks, with a fewer number of nodes, in subjects with more severe cognitive impairment.

Application of speech graph analysis for sorting the groups showed moderate to good classification quality. When selected SGA were combined to the Lawton Index, better classification were obtained, suggesting that the combination of these two

Table 7 | Additional description of verbal fluency and Speech Graph Attributes for the MCI subtypes, and the four groups comparison.

	aMCI		a+mdMCI			<i>p</i>	
	Median	IQR	Median	IQR			
		Q1	Q3	Q1	Q3		
VFE	0	0	0	0	0	1.0000	
VFPR	0.083	0	0.125	0	0	0.1300	
VER	1	0	2	0	1	0.2084	
VFC	11	10	14	11	9	0.0000	
VFTT	13	11	16	11	10	0.0000	
WC	13	11	16	11	10	0.0000	
N	11	11	14	11	9	0.0000	
E	12	10	15	10	9	0.0000	
RE	0	0	0	–	–	0.5682	
PE	0	0	0	0	0	0.7670	
L1	0	–	–	0	0	0.3916	
L2	0	0	0	0	0	0.8658	
L3	0	0	0	–	–	0.0567	
LSC	4	1	7	1	1	0.5115	
ATD	2	1.810	2.095	1.857	1.8	0.1998	
Diameter	9	8	12	9	6	0.0003	
ASP	3.666	3.309	4.666	3.666	2.666	4.333	0.0002
CC	0	0	0	0	0	0.3936	
Density	0.082	0.071	0.090	0.9	0.075	0.109	0.0000

Red values have significance *p* = 0.0083.

simple tools of network measure and functionality can provide to the clinician a good indication of differential diagnosis, except for the contrast between the two MCI subtypes, which spanned a continuum and did not allow the differentiation and classification of the two groups.

The differences prevalent across all groups were in the global attributes of diameter, density and average shortest path (ASP). The results indicate that the networks built by the normal control elderly were more direct, without reoccurrence of words, resulting in a less dense network. Conversely, cognitive impairment corresponded to denser and less direct networks. The density

Table 8 | Pairwise group comparison in the variables with significant difference across the four groups.

	NC × aMCI			NC × a+mdMCI			NC × AD		
	W	z	p	W	z	p	W	z	p
Katz	625	-0.960	0.3371	625	-0.960	0.3371	598.5	-1.388	0.1651
Lawton	548.0	2.581	0.0099	484	3.758	0.0002	325	-6.254	0.0000
MMSE	583.5	-1.046	0.2956	504	-2.597	0.0094	394	4.733	0.0000
VFC	525	-2.186	0.0288	476	-3.140	0.0017	396.5	4.777	0.0000
VFTT	563.5	-1.440	0.1498	501	-2.661	0.0078	410.5	4.514	0.0000
WC	563.5	-1.440	0.1498	501	-2.661	0.0078	410.5	4.514	0.0000
N	526	-2.170	0.0300	475	-3.159	0.0016	394	4.824	0.0000
E	563.5	-1.440	0.1498	501	-2.661	0.0078	410.5	4.514	0.0000
L3	625.0	-0.566	0.5714	612.5	-1.400	0.1614	600	-1.138	0.2552
Diameter	546	-1.778	0.0753	515.5	-2.367	0.0179	425.5	4.239	0.0000
ASP	534.5	-1.994	0.0462	507.5	-2.518	0.0118	414.5	4.433	0.0000
Density	510.5	2.460	0.0139	467.5	3.296	0.0010	388.5	-4.924	0.0000
	aMCI × a+mdMCI			aMCI × AD			a+mdMCI × AD		
Katz	637.5	NaN	NaN	587.5	-2.001	0.0454	587.5	-2.001	0.0454
Lawton	577.5	-1.294	0.1958	352	-5.423	0.0000	350.5	-5.335	0.0000
MMSE	545	1.797	0.0723	416	4.301	0.0000	467.5	3.284	0.0010
VFC	582	1.074	0.2828	471.5	3.334	0.0009	512	2.561	0.0105
VFTT	573.5	1.238	0.2157	466.5	3.418	0.0006	515	2.486	0.0129
WC	573.5	1.238	0.2157	466.5	3.418	0.0006	515	2.486	0.0129
N	574.5	1.223	0.2213	460.5	3.551	0.0004	505	2.692	0.0071
E	573.5	1.238	0.2157	466.5	3.418	0.0006	515	2.486	0.0129
L3	625	0.960	0.3371	587.5	-1.654	0.0981	575	-2.268	0.0233
Diameter	600.5	0.712	0.4764	495.5	2.901	0.0037	549.5	1.884	0.0595
ASP	602.5	0.671	0.5023	502.5	2.755	0.0059	542.5	2.010	0.0444
Density	597	-0.778	0.4367	467	-3.415	0.0006	504.5	-2.700	0.0069

Red values have significance $p = 0.0083$.

differences across the groups were, among all comparisons, the most uniform result, except for the comparison between the two MCI subgroups, which yielded a pattern of continuous performance. The progressive worsening of cognitive performance within the MCI subtypes is consistent in the literature, indicating that a group of subtle deficits underlie the differential diagnosis (Diniz et al., 2007; Radanovic et al., 2009).

Even the groups that did not differ in total number of word repetitions differ in the occurrence of loops of 3 nodes (L3). Nearly all subjects, as expected, managed to avoid recurrences, but 20% of the AD subjects repeated the same word with only two words of interval (e.g., dog-cat-horse-dog). According to Huntley and Howard (2010), subjects with AD already have working memory deficits at the earliest stages of the disease. The impairment in central executive and episodic buffer functions of working memory probably stems from the difficulty of keeping information in mind while keeping the search for new information. These deficits probably explain the repetition of words in verbal fluency tasks with a very small interval.

The results outline a field that needs to be further explored in future studies, involving the density of the networks and the strength between the words in the semantic memory of elderly with pathological aging. The Parallel Distributed Processing Approach of Semantic Cognition predicts that the decrease in

strength of the links between words in a semantic network may allow connections between pairs of words that would not be preferential under normal circumstances (McClelland and Rogers, 2003). Another aspect that deserves further investigation is the absence of difference across the groups in the connectivity attributes (LSC, ATD, and CC). This raises the hypothesis that even very different networks can share a similar structure of local connections, in which a small portion of the words are highly connected with other less connected words, maintaining the integrity of the network's general connection (Bales and Johnson, 2006; De Deyne and Storms, 2008).

Considering the graph analysis performed in this study, build-up in a co-occurrence of the words and based on the temporal link between them, future studies should consider multidimensional scaling and hierarchical clustering analysis. These types of analyses will represent the relation between the variables and combine it into groups, enhancing the results. Future studies should also address the differences between MCI patients and other neurological conditions in which cognitive impairments are quite similar, for example, Temporal Lobe Epilepsy (Holler and Trinka, 2014), as well as the potential association between graph analysis, neuroimaging and other diagnosis instruments. Furthermore, longitudinal studies are also necessary to evaluate whether SGA can help to identify MCI subjects with higher risk of progressing

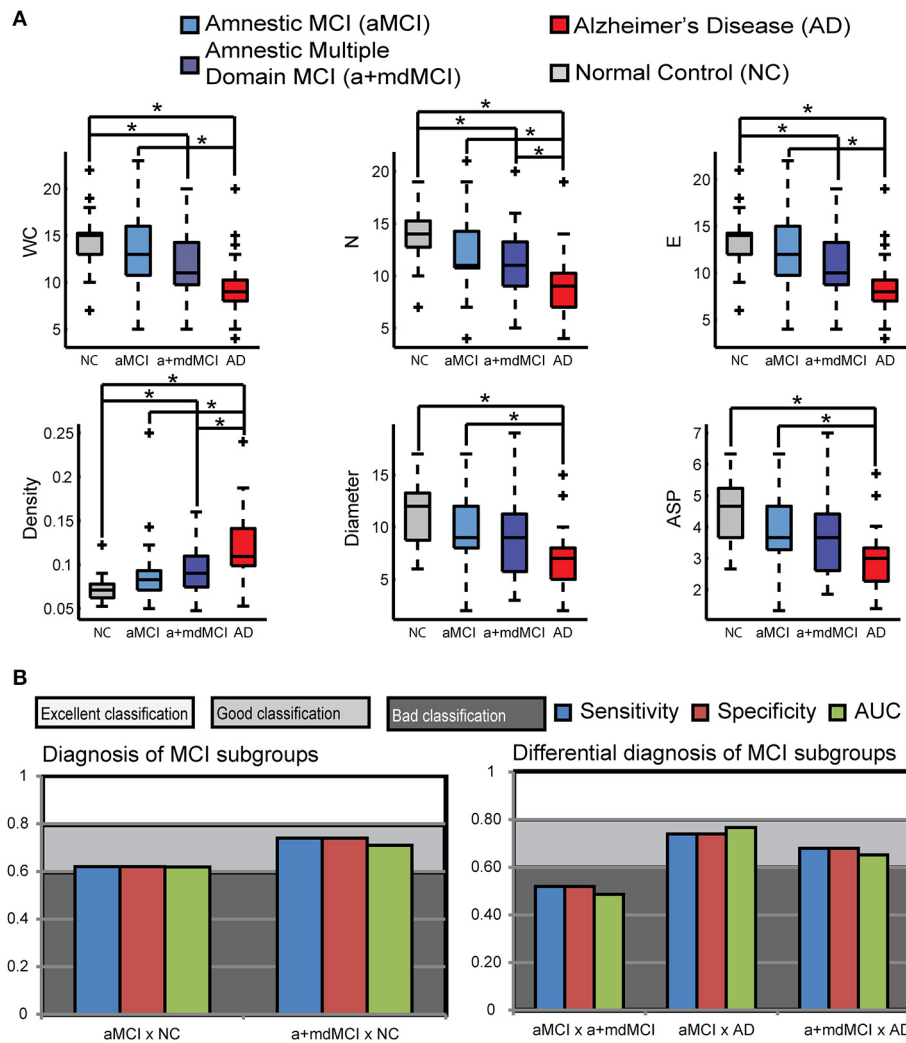


FIGURE 3 | Speech Graph Attributes (SGA) differentiates

psychopathological MCI subgroups. (A) SGA boxplots with significant differences among Alzheimer Disorder (AD), Amnesic Moderate Cognitive Impairment (aMCI), Multiple Domain Moderate Cognitive Impairment (a+mdMCI), and control groups indicated ($N = 25$ per group; Kruskal-Wallis test followed by two-sided Wilcoxon Rank-sum test with Bonferroni correction with $\alpha = 0.0083$). **(B)** Rating quality measured by AUC,

sensitivity and specificity, using SGA correlated with clinical symptoms measured with MMSE and Lawton scales (**Table 4**) (attributes: WC, N, E, Density, Diameter, and ASP). Notice that it is possible to sort the MCI subgroups from the NC or AD groups, but not one from another. Classification quality was considered excellent when AUC was higher than 0.8, good when AUC ranged from 0.6 to 0.8, and poor (not above the chance), when AUC was smaller than 0.6. * $p = 0.0083$.

to Alzheimer's disease. In conclusion, the results suggest that SGA may be a useful tool to help in the differential diagnosis between MCI and AD.

ACKNOWLEDGMENTS

Support obtained from: CNPq Universal 481351/2011-6 and 480053/2013-8, PQ 306604/2012-4 and 308558/2011-1, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), FAPERN/CNPq Pronem 003/2011, FACEPE/CNPq PRONEM APQ-1415-1.05/10, Capes SticAmSud, FAPESP Center for Neuromathematics (grant #2013/ 076990, FAPESP), CBB-APQ-337 00075-09, APQ-01972/12-10 and APQ-02755-10 from FAPEMIG; and 573646/2008-2 from CNPq. Dr. Diniz is supported by grant from the Intramural Research Program

(UFMG) and CNPq (472138/2013-8). The funders had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript. We thank R. Furtado and P. Petrovitch for IT support, A. Karla for administrative help, and D. Koshiyama for library support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/Journal/10.3389/fnagi.2014.00185/abstract>

REFERENCES

Adlam, A.-L. R., Bozeat, S., Arnold, R., Watson, P., and Hodges, J. R. (2006). Semantic knowledge in mild cognitive impairment and mild Alzheimer's disease. *Cortex* 42, 675–684. doi: 10.1016/s0010-9452(08)70404-0

- Amieva, H., Jacqmin-Gadda, H., Orgogozo, J.-M., Le Carret, N., Helmer, C., Letenneur, L., et al. (2005). The 9 year cognitive decline before dementia of the Alzheimer type: a prospective population-based study. *Brain* 128, 1093–1101. doi: 10.1093/brain/awh451
- Aretouli, E., Okonkwo, C. O., Samek, J., and Brandt, J. (2011). The fate of the 0.5s: predictors of 2-year outcome in mild cognitive impairment. *J. Int. Neuropsychol. Soc.* 17, 277–288. doi: 10.1017/s1355617710001621
- Bales, M. E., and Johnson, S. B. (2006). Graph theoretic modeling of large-scale semantic networks. *J. Biomed. Inform.* 39, 451–464. doi: 10.1016/j.jbi.2005.10.007
- Brucki, S. M. D., Malheiros, S. M. F., Okamoto, I. H., and Bertolucci, P. H. F. (1997). Dados normativos para o teste de fluência verbal categoria animais em nosso meio. *Arq. Neuropsiquiatr.* 55, 56–61.
- Butts, C. T. (2009). Revisiting the foundations of network analysis. *Science* 325, 414–416. doi: 10.1126/science.1171022
- Craik, F. I. M., and Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends Cog. Sci.* 10, 131–138. doi: 10.1016/j.tics.2006.01.007
- De Deyne, S., and Storms, G. (2008). Word associations: network and semantic properties. *Behav. Res. Methods* 40, 213–231.
- de Paula, J. J., Bertola, L., Ávila, R. T., Moreira, L., Coutinho, G., Moraes, E. N., et al. (2013). Clinical applicability and cutoff values for an unstructured neuropsychological assessment protocol for older adults with low formal education. *PLoS ONE* 8:E73167. doi: 10.1371/journal.pone.007316
- Diniz, B. S. O., Yassuda, M. S., Nunes, P. V., Radanovic, M., and Forlenza, O. V. (2007). Mini-mental state examination performance in mild cognitive impairment subtypes. *Int. Psychogeriatr.* 19, 647–656. doi: 10.1590/S1516-44462008000400003
- Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., and Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: a new approach. *Brain Lang.* 93, 79–94. doi: 10.1016/j.bandl.2004.08.003
- Goni, J., Arrondo, G., Sepulcre, J., Martincorena, I., Mendizabal, N. V., Corominas-Murtra, B., et al. (2010). The semantic organization of the animal category: evidence from semantic verbal fluency and network theory. *Cogn. Process.* 12, 183–196. doi: 10.1007/s10339-010-0372-x
- Griffiths, T. L., Steyvers, M., and Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychol. Rev.* 114, 211–244. doi: 10.1037/0033-295X.114.2.211
- Henry, J. D., Crawford, J. R., and Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia* 42, 1212–1222. doi: 10.1016/j.neuropsychologia.2004.02.001
- Hodges, J. R., Erzinçlioglu, S., and Patterson, K. (2006). Evolution of cognitive deficits and conversion to dementia in patients with mild cognitive impairment: a very-long-term follow-up study. *Dement. Geriatr. Cogn. Disord.* 21, 380–391. doi: 10.1159/000092534
- Holler, Y., and Trinka, E. (2014). What do temporal lobe epilepsy and progressive mild cognitive impairment have in common? *Front. Syst. Neurosci.* 8:58. doi: 10.3389/fnsys.2014.00058
- Huntley, J. D., and Howard, R. J. (2010). Working memory in early Alzheimer's disease: a neuropsychological review. *Int. J. Geriatr. Psychiatry* 25, 121–132. doi: 10.1002/gps.2314
- Kotsiantis, S. B. (2007). "Supervised machine learning: a review of classification techniques." in *Emerging Artificial Intelligence Applications in Computer Engineering: Real Word Ai Systems with Applications* (Amsterdam: IOS Press), 3–24.
- Lerner, A. J., Ogrocki, P. K., and Thomas, P. T. (2009). Network graph analysis of category fluency testing. *Cog. Behav. Neurol.* 22, 45–52. doi: 10.1097/wnn.0b013e318192ccaf
- Lezak, M., Howieson, D. B., and Loring, D. W. (2004). *Neuropsychological Assessment, 4th Edn.* New York, NY: Oxford University Press.
- McClelland, J. L., and Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nat. Rev. Neurosci.* 4, 310–323. doi: 10.1038/nrn1076
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's. *Neurology* 34, 939–944. doi: 10.1212/WNL.34.7.939
- Mota, M. B., Furtado, R., Maia, P. P. C., Copelli, M., and Ribeiro, S. (2014). Graph analysis of dream reports is especially informative about psychosis. *Sci. Rep.* 4:3691. doi: 10.1038/srep03691
- Mota, N. B., Vasconcelos, N. A. P., Lemos, N., Pieretti, A. C., Kinouchi, O., Cecchi, G. A., et al. (2012). Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS ONE* 7:e34928. doi: 10.1371/journal.pone.0034928
- Nickles, L. (2001). "Spoken word production," in *What Deficits Reveal about the Human Mind/Brain: A Handbook of Cognitive Neuropsychology*, ed B. Rapp (Philadelphia, PA: Psychology Press), 291–320.
- Nutter-Upham, K. E., Saykin, A. J., Rabin, L. A., Roth, R. M., Wishart, H. A., Pare, N., et al. (2008). Verbal fluency performance in amnesic MCI and older adults with cognitive complaints. *Achiev. Clin. Neuropsychol.* 23, 229–241. doi: 10.1016/j.acn.2008.01.005
- Patterson, K., Nestor, P. J., and Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–988. doi: 10.1038/nrn2277
- Radanovic, M., Diniz, B. S., Mirandez, R. M., Novaretti, T. M. S., Flacks, M. K., Yassuda, M. S., et al. (2009). Verbal fluency in the detection of mild cognitive impairment and Alzheimer's disease among Brazilian portuguese speakers: the influence of education. *Int. Psychogeriatr.* 21, 1–7. doi: 10.1017/S1041610209990639
- Salmon, D. P., Thomas, R. G., Pay, M. M., Booth, A., Hofstetter, C. R., Thal, L. J., et al. (2002). Alzheimer's disease can be accurately diagnosed in very mildly impaired individuals. *Neurology* 59, 1022–1028. doi: 10.1212/wnl.59.7.1022
- Saxton, J., Lopez, O. L., Ratcliff, G., Dulberg, C., Fried, L. P., Carlson, M. C., et al. (2004). Preclinical Alzheimer disease: neuropsychological test performance 1.5 to 8 years prior to onset. *Neurology* 63, 2341–2347. doi: 10.1017/S1041610208007631
- Singh, M., and Provan, G. M. (1995). A comparison of induction algorithms for selective and non-selective Bayesian classifiers. *Mach. Learn. Proc.* 1995, 497–505.
- Strauss, E., Sherman, E. M. S., and Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* Oxford, UK: Oxford University Press.
- Taler, V., and Phillips, N. A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *J. Clin. Exp. Neuropsychol.* 30, 501–556. doi: 10.1080/13803390701550128
- Unsworth, N., Spillers, G. J., and Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Q. J. Exp. Psychol.* 64, 447–466. doi: 10.1080/17470218.2010.505292
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., et al. (2004). Mild Cognitive impairment—beyond controversies, towards a consensus: reports of the international working group on mild cognitive impairment. *J. Int. Med.* 256, 240–246. doi: 10.1111/j.1365-2796.2004.01380.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 March 2014; accepted: 09 July 2014; published online: 29 July 2014.

Citation: Bertola L, Mota NB, Copelli M, Rivero T, Diniz BS, Romano-Silva MA, Ribeiro S and Malloy-Diniz LF (2014) Graph analysis of verbal fluency test discriminate between patients with Alzheimer's disease, mild cognitive impairment and normal elderly controls. *Front. Aging Neurosci.* 6:185. doi: 10.3389/fnagi.2014.00185 This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Bertola, Mota, Copelli, Rivero, Diniz, Romano-Silva, Ribeiro and Malloy-Diniz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.