Hindawi
Complexity
Volume 2021, Article ID 6398407, 13 pages
https://doi.org/10.1155/2021/6398407

Research Article

Navigating Concepts in the Human Mind Unravels the Latent Geometry of Its Semantic Space

Barbara Benigni[1,2] **Monica Dallabona,**[3] **Elena Bravi,**[3] **Stefano Merler**[4] **and Manlio De Domenico**[2]

[1]*Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 9, 38123 Povo, Trento, Italy*
[2]*CoMuNe Lab, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo, Trento, Italy*
[3]*Department of Mental Health, Division of Psychology, Azienda Provinciale per i Servizi Sanitari, Viale Verona 38123, Trento, Italy*
[4]*DPCS, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo, Trento, Italy*

Correspondence should be addressed to Barbara Benigni; bbenigni@fbk.eu

Received 11 June 2021; Accepted 5 November 2021; Published 13 December 2021

Academic Editor: Hang-Hyun Jo

Mechanisms for retrieving basic knowledge in the human mind are still unknown. Exploration is usually represented by cognitive units, i.e., concepts, linked together by associative relationships forming semantic networks. However, understanding how humans navigate such networks remains elusive, because the underlying topology of concepts cannot be observed directly, and only functional representations are accessible. Here, we overcome those limitations and show that the hypothesis of an underlying, latent geometry characterizing the human mind is plausible. We characterize this geometry by means of adequate descriptors for exploring and navigating dynamics, demonstrating that they can capture the differences between healthy subjects and patients at different stages of dementia. Our results provide the first fundamental step to develop a new unifying conceptual and computational framework that can be used to support the assessment of neurodegenerative diseases from language and semantic memory retrieval tasks, as well as helping develop targeted nonpharmacological therapies to maintain residual cognitive capacity.

1. Introduction

The retrieval of basic knowledge from memory, known as semantic memory [1], has long been the focus of a lively debate across multiple research fields. Such a debate mainly divides between two schools of thought, i.e., the one of semantic space and the one of semantic networks. According to the branch of semantic space, the search is a key cognitive feature that operates similarly across different scales and contexts [2]. In many domains, the search at different scale (from the search for an object in a bag to the search for a disease cure) always requires to manage the trade-off between exploiting what is known and exploring what is unknown [2]. In this sense, the internal search of memory retrieval exhibits similar characteristics to the external search in physical space [3]. According to the theory of optimal foraging [3], the process of retrieving concepts from memory is dynamically similar to the one performed by animals when searching for food between patches of their environment [4]. This mental dynamical process mediates between local exploitation of clusters of information and global exploration of such cluster, pursuing a sort of semantic foraging [3]. In accordance with the marginal value theorem [5], the semantic memory search is considered optimal if the subject, as the animal does in the optimal foraging, leaves a given cluster of information when the benefit of local exploitation falls into the level of the expected benefit of changing cluster and searching elsewhere [3]. In the clinical field, the intuition that patients cognitively organized the semantic access around semantic clusters following a clustering and switching pattern in search has been widely used to investigate the semantic retrieval [6] and the semantic impairment [7]. However, in this method, clusters are based on hand-made classification according to taxonomies: limitations can be partially overcome by taking advantage of distributional semantics to define the clustering and chaining of concepts during a semantic memory retrieval task [8]. In summary, according to the semantic space school of thought, the modelling of searching in semantic memory needs two main ingredients: (a) a structural representation of the search space (hand-coded or statistically derived) and (b) a model of the search process (e.g., local to global transitions) [4]. However, there is still no clear definition of what a patch is and how to define it in memory [4]. Concurrently, another school of thought, the one of semantic network, demonstrated that the same results obtained with the optimal foraging in semantic space could emerge from a random walk exploring a semantic network [9, 10]. Instead of the clustering and switching processes, this network approach was postulating a simpler and single process of exploration on a network of concepts. According to this approach, the navigation of concepts is represented by associative semantic networks [11], fostering the idea that concepts are cognitive units, each represented as a node linked to associated elements [12–15]. A typical issue often leveled against the use of semantic networks is that they might end up explaining, or predicting, memory retrieval by leveraging on models built from similar behaviors, for example, when modelling semantic networks from free associations data to explain semantic fluency tasks [16]. Progress in building such networks from fluency data has been made [17]; however, there is still no consensus about the most appropriate way to construct semantic networks [18]. Nevertheless, the semantic network approach has been widely used in the clinical field for the assessment of psychosis [19], Alzheimer's disease [20–22], and in cognitive science, for example, to investigate the levels of creativity [23] and the openness to experience in the human beings [24]. Over the past two decades, vector-space models of words meaning as high-dimensional numerical vectors have become serious contenders of semantic representation [25], for example, when studying human psycholinguistic tasks [26] or when exploring the semantic verbal fluency in mild cognitive disorder [8]. Powerful tools involving this kind of spatial representation of words are the so-called word embeddings, bridging distributional semantics and natural language processing, which map words into vectors in a multidimensional space [27]. The underling idea of this approach can be summarized with the words of the English linguist J. R. Firth, "a word is characterized by the company it keeps" that, from a mathematical perspective, means that the closer the words in the multidimensional space, the closer their meaning in the vocabulary.

Language and semantic memory retrieval tasks are crucial in the identification of neurodegenerative diseases [28–30] and are usually employed in different neuropsychological tests. Among these tests, the categorical semantic verbal fluencies (SVF) play an important role in the assessment of dementia and Alzheimer's disease in particular [6, 31]. Here, patients are asked to pronounce as many words as possible, belonging to a certain category, within a given time interval. Patients' performance is successively evaluated, in particular by counting the amount of words pronounced [32] or their response times [33]. To investigate semantic retrieval, further approaches, based on the intuition that patients cognitively organized the semantic access around semantic clusters, have been widely used [6]. Recently, evidence of semantic maps tiling human cerebral cortex has been provided from fMRI data, probing the existence of semantic selectivity in brain areas [34] and further strengthening the insight that the language can be organized on a topological space, i.e., on a manifold. Nevertheless, a clear understanding of the mechanisms behind the navigation of semantic memory still eludes us [35, 36].

The aim of this work is to provide a data-driven insight on why the hypothesis of a spatial representation, i.e., of an underlying, latent geometry characterizing the human mind, is plausible. Relying on clinical data of semantic verbal fluency test from 215 subjects and leveraging on word-embeddings tool, we aim at defining suitable metrics to indirectly explore this possible, latent geometry. Our work arises within the debate between semantic space and semantic network representations, and it examines the exploration process, integrating both perspectives, space and networks, according to an approach that builds upon the geometry at the concept scale and culminates with diagnosis-based semantic networks, passing through the mesoscale organization of concepts (clustering). Remarkably, our framework allows us to gain new insights into the organization of concepts in the human mind and shed some light on why some existing approaches were successful. In fact, the mechanisms behind the retrieval of basic knowledge, known as semantic memory [1], still remain fundamentally unknown [37]. Here, we fill this gap by (i) hypothesizing the existence of an underlying geometry, which governs the exploration of concepts in the human mind, and (ii) demonstrating that such a geometry can discriminate between healthy subjects and patients at different stages of dementia. Our hypotheses are based on the assumption that if a common latent geometry underlying the mental navigation of concepts existed, then subjects with semantic retrieval deficits should show some distortion in such a navigation on top of this geometry. Here, we observe how different population of subjects, in terms of semantic impairment, differently navigates the same geometry by means of suitable metrics characterizing their explorations. If our hypothesis is reasonable, we predict to see significant differences in metrics computed from different diagnoses.

Our study is based on the analysis of semantic verbal fluencies (SVF) data, belonging to animal category, from 92 patients suffering of dementia (DEM, $M = 40\%$, $F = 60\%$, age $= 75 \pm 7$, yrs of education $= 9 \pm 4$), 93 patients suffering of Mild Cognitive Impairment, a precursor of Alzheimer's disease (MCI, $M = 48\%$, $F = 52\%$, age $= 77 \pm 6$, yrs of education $= 9 \pm 4$), and 30 healthy controls (CTR, $M = 60\%$, $F = 40\%$, age $= 32 \pm 7$, yrs of education $= 17 \pm 0.40$). During the semantic verbal fluency test, each individual is asked to report all words he/she can remember belonging to category of animals, within a time interval of 60 seconds. Each spoken word is annotated by the neuropsychologist who is testing the patient. No clues nor incentives are given to the subjects during the tests, and any repetitions are not marked. The SVF test is a significant test for the assessment of dementia diagnosis [31]. Generally, the semantic impairment is more severe in patients with dementia than MCI patients. The rationale for looking at the semantic fluencies of these two populations aims at testing our guess that if a latent geometry existed, a different severity in semantic memory retrieval impairment should be reflected in a different way of navigating concepts on such a plausible, latent geometry. MCI subjects have an increased risk of conversion to Alzheimer's disease (and dementia in general). Possibly, testing the differences between the exploration of concepts of these two populations and a group of control by means of spatial metrics might be relevant to get insight into how the navigation of different category of subjects is performed from a data-driven perspective. Instead of focusing only on statistical descriptors of language, e.g., word frequency or vocabulary size, in disease, we considered also in which sequence the words have been provided; this information is crucial, because it allows to map the navigation of concepts in the underlying semantic space.

To characterize the navigability of this space in terms of concepts visited in such unknown, possibly multidimensional space, we first had to build a plausible geometric proxy (illustrative representation in Figure 1).

To this aim, we used three distinct word embeddings obtained from the Italian language, namely, Italian Word Embeddings, trained on the Italian Wikipedia [38], itWac, constructed from the Web limiting the crawl to the .it domain and using medium-frequency words from the Repubblica corpus and basic Italian vocabulary lists as seeds [39], and Twitter, trained on 46.935.207 tweets [39]; all the word embeddings were generated with the popular word representation models, word2vec [27]. By choosing three different word embeddings, we are able to evaluate the robustness of our metrics in geometries coming from different sources, i.e., a website (Wikipedia), a social network (Twitter), and a newspaper (La Repubblica). In the following, we will refer to word embeddings, semantic spaces, or geometries interchangeably. Here, the term geometry is justified by the fact that we leveraged on word-embeddings, powerful tools that encode the semantic relation between the words as a geometric relationship between vectors in multidimensional space. Word embeddings are built from data corresponding to humans-written documents (in our case, Wikipedia, Twitter, and La Repubblica, an Italian newspaper), which are then embedded in a multidimensional space according to the hypothesis of the distributional semantics. This hypothesis defines the semantics similarity in terms of vector similarity; i.e., the closer the meaning in the vocabulary, the closer the points representing the words in the word embeddings (encoded by a vector of coordinates in a multidimensional space). In this sense, by embedding the words pronounced by a sample of subjects into a word embedding, which is a coordinate space by design, we can study the mental navigation of such subjects on a geometry of concepts. For each group of subjects, we, therefore, have three independent semantic spaces; each one is used to characterize the local exploration and the overall navigation of the semantic geometry. More specifically, we introduce five different descriptors for this purpose, in order to identify the effects of the underlying geometry, if any. At the smallest scale, i.e., the one of single concepts, geometry is probed in terms of:

(1) Maximum jump Max$J$, i.e., the maximum distance, in the word embeddings, between two consequent words pronounced during the test, it defines the maximum instantaneous capacity to change context;
(2) Diameter of exploration DOE, i.e., the maximum distance, in the word embeddings, between the words pronounced during the test, whatever the order, it defines the maximum capacity to change context in the whole test duration. To be consistent, we call this metric DOE, when it is computed with the Euclidean distance, and amplitude of exploration (AOE), when it is computed with the cosine distance;
(3) Density of exploration ρ_w, it corresponds to the total amount of animal words potentially available in the hypersphere built from the exploration radius R, half the DOE, and which has as its center the centroid C of spoken words in the word embeddings. It returns a measure of density in the volume of words explored by the subject. Specifically, it defines the density of pertaining words (i.e., belonging to the category of animal) in the area explored by the subject in the geometry;
(4) Distance d, it is the total distance covered during the test; it quantifies the magnitude of the overall exploration;
(5) And farness far, i.e., the average distance of the words pronounced, it defines the ability to go far with a certain number of jumps.

For mathematical details about each descriptor, we refer to section Methods, while their significance in discriminating the three groups of subjects is evaluated by means of Kolmogorov–Smirnov statistical tests and t-tests.

Afterwards, in line with the idea that the semantic access is cognitively organized around semantic clusters [3, 6], we probed the mesoscale organization of concepts by performing a semisupervised clustering algorithm in the three geometries. Accordingly, we define the explorative potential of the navigation for each category as the total number of visited clusters and as the total number of words included in the visited clusters. This descriptor is a proxy for the cognitive effort spendable during the navigation, and it defines the total amount of clusters/words, potentially visitable/retrievable during the test. Clusters are then given as input to a hierarchical clustering algorithm, which provides the spatial hierarchy of such clusters based on their relative distance. By comparing the distances between the visited clusters, we are able to evaluate the existence of a hierarchy in the way subjects explore concepts (technical details about clustering and explorative potential can be found in section Methods).

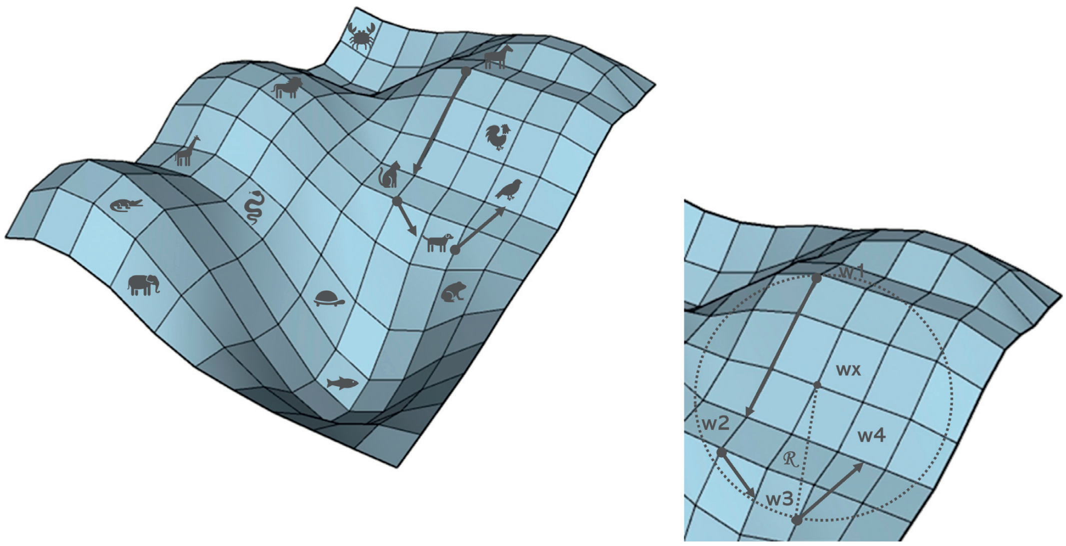

Figure 1: Conceptual representation of semantic space. On the left, navigation of concepts on the semantic space, arrows, defines the sequence of words. On the right, a zoom outlining the navigation, w_n, is the concepts, w_x their centroid, and R is the radius of exploration. This figure has been generated using the 3D Design Software SketchUp 2020.

Taking inspiration from the process of clustering and switching when retrieving concepts from memory, network scientists provided a new kind of random walk over a graph as a Markov process, i.e., the switcher random walk [40], to generalize the exploration task on a network. In this vein, and by following the assumption of the semantic network navigated by a random walk [10], we finally tested the navigation of concepts by means of their Markov representation, to probe the possible alterations of mental pathways emerging from the exploration of concepts in patients with dementia. Mathematically, this corresponds to defining the transition probability from one state (i.e., cluster) to another, regardless of the previously visited states. Operationally, we build three Markov chains, one for each group, i.e. the two diagnosis and the healthy control, considering all the clusters visited by each group as the states of the Markov chain of that group, and setting the transition probabilities equal to the relative transition frequencies from one state to another in each group. Each Markov chain is characterized by the steady state distribution and the mean first passage time matrix. In the following, we provide the intuitions on how to interpret these two descriptors for each network of concepts. The intuitions behind the steady state distribution and the mean first passage time (MFPT) matrix are given by the purpose to investigate and to characterize the search process pursued by each diagnosis on its network of concepts. Being the mathematical model of the network of concepts assumed as a Markov chain, the steady state distribution and the MFPT are the key descriptors to investigate such a navigation dynamic. We assume that if it is true that different diagnoses explore the network of concepts in different way, the steady state distributions and the MFPT should highlight these differences. In fact, the steady state distribution defines the unique distribution to which the exploration converges as the number of transitions increases, regardless of a Markov chain's initial state. Here, the steady state is a vector, computed for each category, representing the probabilities to be in each of the cluster of words visited by that category, after a sufficient amount of time. It is to be noticed that each subject has one minute to complete the SVF test, but practically no patient uses it all because he finishes the words before the one minute ends. In this sense, one minute is enough to reach a regime situation, which is mathematically represented through the steady state distribution. However, it would be experimentally impossible to test a subject for an infinite amount of time. Our intention is to compare metrics that uniquely identify the pattern of exploration, as given by steady state and MFPT, for each category of subject; in this way, we can detect any possible differences between such patterns. For what concerns the intuition behind the MFPT matrix, it encodes the mean amount of time required to go from one state i to another state j of a Markov chain. In our case, the MFPT matrices encode the mean number of transitions to go from one cluster of words to another. Specifically, we define a MFPT matrix for each diagnosis. The entries of such a MFPT matrix answer the question: starting form one cluster of words i, how long does it take, on average, for this specific category to reach a specific cluster of words j for the first time? In this sense, the MFPT matrix characterizes the exploration dynamic of each diagnosis since it returns an average measure of the time spent to navigate the underlying network of concepts. In summary, MFPT matrix defines the average number of steps needed to reach a certain state from another for the first time. This idea, redefined on the network of concepts, corresponds to the average time required for each diagnosis to pass from one cluster to another for the first time and then enable us to measure the time needed to travel for the first time a certain mental link connecting two groups of concepts. Taken in isolation, the steady state distribution (ss) and the MFPT can give us an insight of how the exploration of each diagnosis evolves on the network and over time. For example, they provide us how heterogeneously the clusters will be explored after a sufficient number of transitions (ss) and how much time it takes before a cluster is visited for the first time (MFPT). To investigate possible differences in the dynamic of exploration between the diagnoses and the healthy control group, these descriptors are then compared by means of similarity measures between the three groups, i.e., Pearson correlation, Spearman's correlation, Euclidean norm, Frobenius norm, and covariance (for mathematical details about Markov chains, we refer the reader to section Methods).

2. Results

2.1. Geometry. Overall, the metrics defined to characterize the local exploration prove to be suitable for discriminating between healthy and nonhealthy subjects in all the three spaces. The results for the three semantic space are shown in Figure 2. Specifically, according to the results of t-tests all the metrics in all geometries, except for far in twitter geometry, are able to discriminate between healthy and nonhealthy subjects, all having p values ≤ 0.0104 (see Table 2 of Supplementary Materials for detailed results of t-tests). Also, according to the results of Kolmogorov–Smirnov statistical test, all the metrics, except for Max_j in itWaC geometry and far in twitter geometry, reveal to be able to discern between healthy and nonhealthy subjects all having p values ≤ 0.0304 (see Table 1 of Supplementary Materials for detailed results of Kolmogorov–Smirnov statistical test).

Remarkably, the distance d is always significant not only in discerning between healthy and nonhealthy subjects, but also between different stages of dementia according to both Kolmogorov–Smirnov and t-test (p values of KS test in all the geometry are ≤ 0.0209, p values of t-test in all the geometry are ≤ 0.0076). Interestingly, in the Wikipedia geometry, KS test highlights that all the metrics turn out to be significant (all having p values ≤ 0.0233), except for far and Max_j, in distinguishing between all the three categories, i.e., healthy controls and the two stages of dementia MCI and DEM. Also for the t-tests, the metrics turn out to be significant (all having p values ≤ 0.017867), except for far, in separating the three categories. Results on local exploration can be summarized as follows:

(i) All the metrics can be used in all the geometries to separate between healthy and nonhealthy, except for Max_j in itWaC geometry and farin twitter geometry;

(ii) All the metrics should be used only in the Wikipedia geometry, excluding the far and Max_j, to discriminate between the three categories (DEM, MCI, and healthy);

(iii) The distance d metric is robust in separating all the three categories across the three word embeddings and should be used when considering the itWaC and the Twitter geometry to discriminate between different stages of dementia.

Detailed results of Kolmogorov–Smirnov statistical tests and t-test for each metric are reported in Tables 1 and 2 of Supplementary Materials.

2.2. Hierarchy. The explorative potential is able to discriminate between healthy and nonhealthy subjects according to both KS tests and t-tests (p values $\leq$ 0.002 6, see Tables 3 and 4 of Supplementary Materials for detailed results), strengthening what we found at the local scale.

Figure 3 shows the tanglegrams for the pair MCI-DEM for the three semantic spaces; it shows as well the values of Baker's Gamma correlation [41] compared with the null models for all the pairs. What is clear in this analysis is the strong correlation between MCI and DEM in the hierarchy through which the concepts are explored, as evidenced by the values of Baker's correlation, equal to 0.88 in itWac, 0.97 in Twitter, and 0.73 in Wikipedia, and as validated by the null model, in contrast with the correlation for all other pairs. This is notably remarkable when compared with the values of correlation between the stages of dementia and healthy controls, which, instead, are always close to zero in the three geometries (two-dash lines in Figure 3).

2.3. Networks. The Markov chains modelling the exploration of concepts are displayed in Figure 4 and can be considered as a proxy of the semantic networks for each group of subjects. The numbers of nodes of such networks vary within the same group according to the spaces because of clustering mapping (itWaC: CTR 19, MCI 13, DEM 11; Twitter: CTR 16, MCI 12, DEM 9; Wikipedia: CTR 32, MCI 25, DEM 28). Overall, nonhealthy patients explore a smaller portion of the semantic nodes with respect to healthy control. It is to be noticed that, for the geometries of itWaC and Twitter, there is a progressive decrease in the number of visited nodes going from CTR to MCI and from MCI to DEM. Not all the considered correlation measures between the steady states and the mean first passage time matrices agree in ranking the similarity between the analyzed groups, and only some specific combinations of geometry-correlation measure highlight higher correlation for the couple MCI-DEM. In particular, this is true for the values of Spearman correlation in itWaC and Twitter geometries and for the ones of Pearson correlation in Wikipedia (specific results are reported in Tables 5–7 of the Supplementary Materials). Finally, the higher correlation between MCI and DEM is confirmed in mean first passage time matrices, where the values of Pearson correlation are the highest for the couple MCI-DEM, while the Euclidean norm of their difference is the smallest, evidence of similarity between the two groups in all the geometries (specific values reported in Tables 8–10 of Supplementary Materials).

3. Discussion

In this work, we investigated the assumption related to the semantic space by testing how plausible is the hypothesis of a

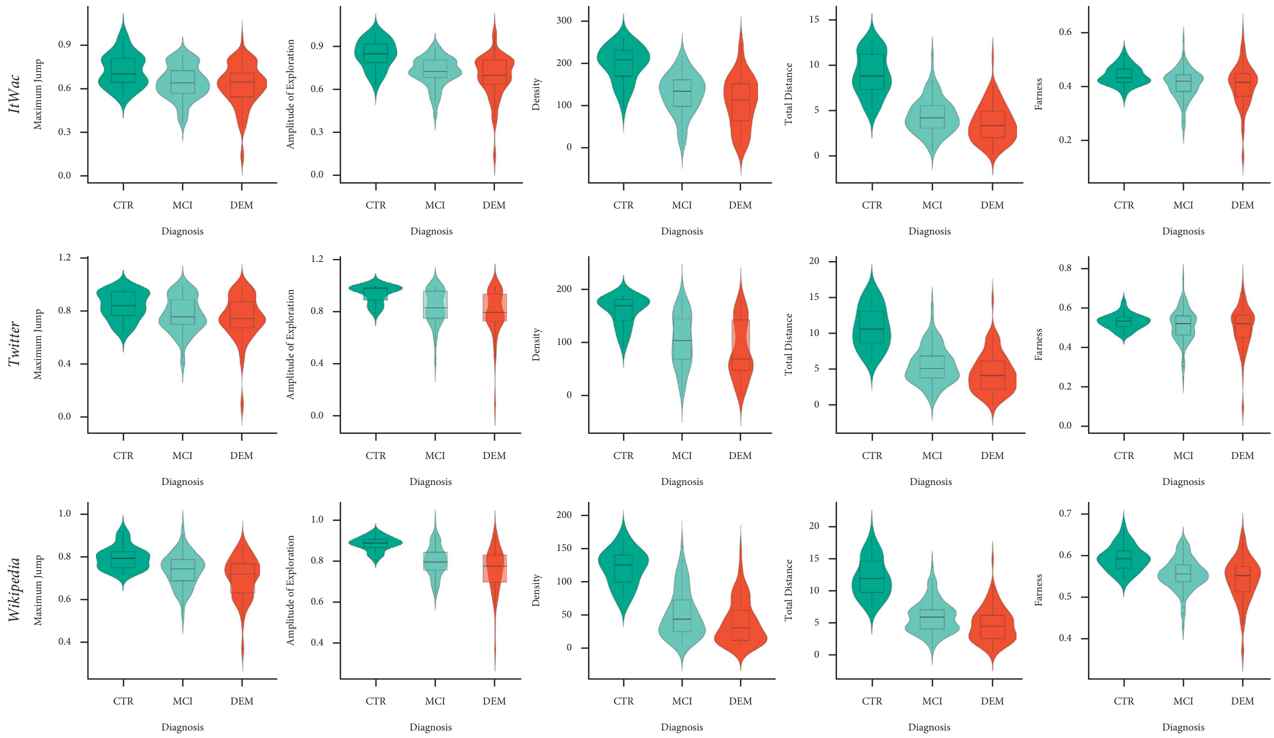

Figure 2: Geometry. Descriptors of local exploration of concepts for CTR, MCI, and DEM. The distributions of the five local descriptors are represented by boxplots within violin plots for each geometry and for each group of subjects. All the metrics are computed using the cosine distance. This figure has been generated using the publicly available R software, version 3.6.3.

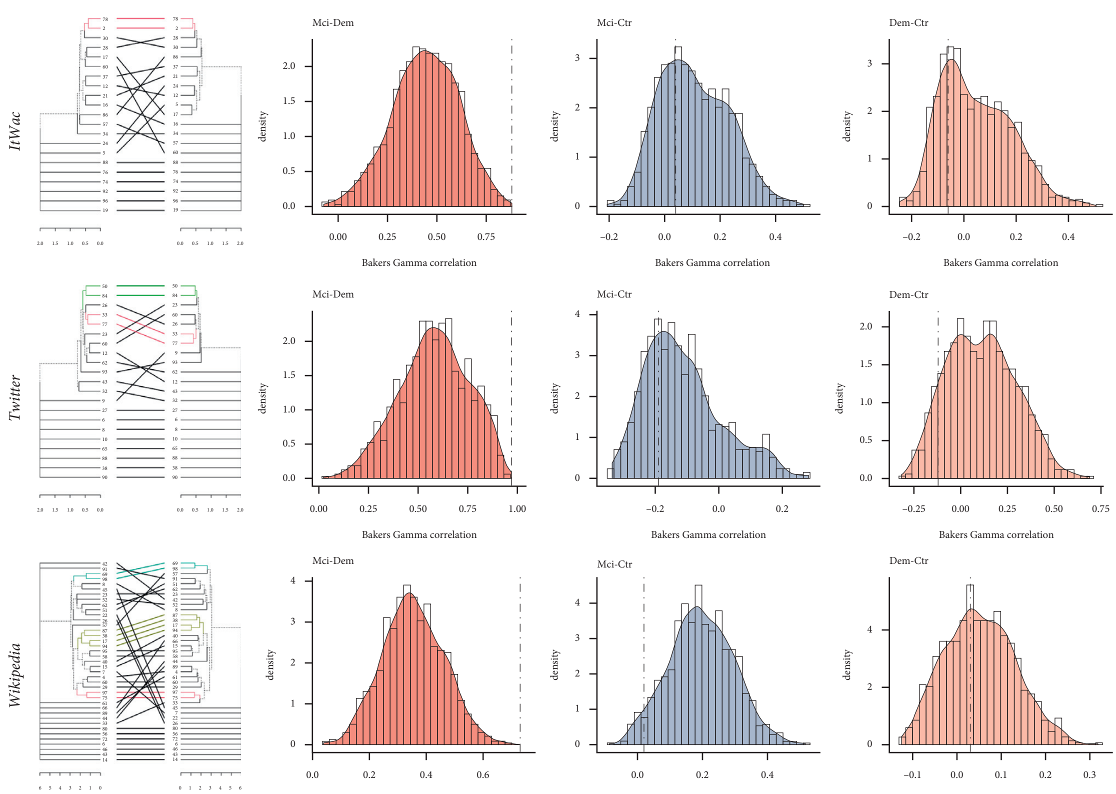

Figure 3: Hierarchy in exploration of concepts. Tanglegrams of the pairs MCI-DEM with null models of Baker's Gamma correlation for all the pairs, for each semantic space. Colored density functions represent Baker's Gamma correlation of the null models, while two-dash vertical intercept is the real value of correlation. This figure has been generated using the publicly available R software, version 3.6.3.

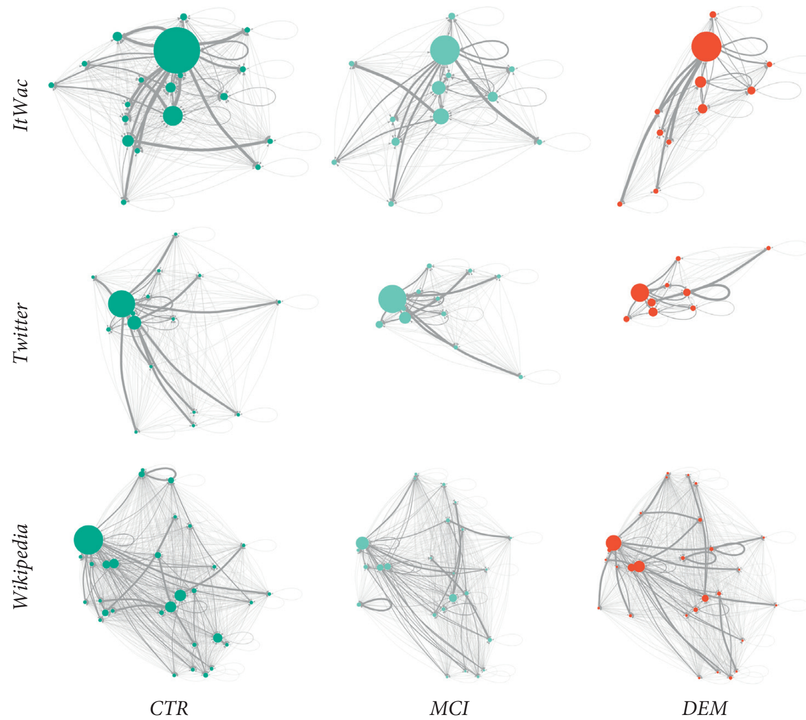

Figure 4: Networks. Mapping mental pathways emerging from the navigation of concepts in healthy controls (CTR), mild cognitive impairment subjects (MCI), and patients with dementia (DEM). Networks of concepts as reconstructed from semantic verbal fluency tests, for the three semantic spaces. Colored nodes encode clusters of concepts reported by patients while performing the test where they are asked to report words belonging to animal category. The size of nodes is proportional to the nodes' strength, while the thickness of the edges is proportional to their weight. This figure has been generated using the publicly available R software, version 3.6.3.

latent geometry underlying exploration of concepts in human mind, and whether this geometry can be used to discriminate between healthy and nonhealthy subjects. Our hypothesis is tested against different types of navigation, i.e., the one coming from healthy subjects and the one coming from subjects with deficit in semantic memory retrieval task according to a prior clinical evaluation. By means of suitable metrics characterizing the spatial navigation of concepts on three distinct word embeddings, we have demonstrated why is being plausible, that the mental navigation process takes place on a latent geometry, understood as an organized manifold of lexical information, by relying on data coming from 215 semantic verbal fluency tests. In terms of this, the geometry of the word embeddings acted as a proxy of a potential geometry of the human mind, intended as the setting where the information is somehow organized, when navigating the concepts.

We examined the exploration process, integrating two main perspectives, space and networks, according to an approach that builds upon the geometry at the concept scale and culminates with diagnosis-based semantic networks, passing through the mesoscale organization of concepts. On the one hand, semantic networks do not give satisfactory strength of evidence in discerning between the groups of subjects considered (CTR, MCI, and DEM). In fact, results vary according to word embeddings and correlation measures, consequently proving to be an indicative but not definitive approach. On the other hand, the geometric approach gives significant results in revealing differences between healthy and nonhealthy subjects through local descriptors, and in highlighting the similarities between Mild Cognitive Impairment and patients with dementia through hierarchy. It is to be noticed that the metrics distance d is always significant not only in discerning between healthy and nonhealthy subjects, but also between different stages of dementia in all the three geometry. Intriguingly, in the Wikipedia geometry, all the metrics, except for farness and maximum jump, are able to separate all the three categories. In short, if we had to choose one metric that can separate the three categories (DEM, MCI, and healthy controls), whatever the geometry is, we would choose the distance d. Otherwise, we had to choose a semantic space that can capture the differences between the three categories in all the considered metrics (excluding the farness and the maximum jump), and we would declare the Wikipedia geometry as the chosen one for this task. Finally, it is always possible to distinguish between healthy and nonhealthy in all the geometries whatever the metric is, except for Max_j in itWaC geometry and far in twitter geometry (this is demonstrated by the results of Kolmogorov–Smirnov tests and t-tests, reported in table from 1 to 4 in the Supplementary Materials). Our results suggest how the metrics, coupled with word embedding, should be chosen according to the purpose (i.e., discriminate between healthy and non-healthy and/or discriminate between all the three considered categories, DEM, MCI, and healthy). It is worth noting that the

Wikipedia word embedding is a multidimensional space of 300 dimensions, that is, more than double compared to the other word embeddings used in this study (itWaC and Twitter), which have 128 dimensions. This means that, to some extent, the Wikipedia geometry contains more information encoded in the relationship between words. Thus, it could be possible that all metrics computed in the Wikipedia geometry can discriminate between all the three categories precisely because of this higher information stored in this word embedding.

We conclude that the geometric framework is an effective and robust approach to investigate the semantic memory retrieval and to assess its abnormal navigation in patients at different stages of dementia. For this reason, our metrics could be used in support of the clinical assessment as a data-driven tool for confirming, and not yet predicting, the diagnosis. This would help planning the longitudinal referral, for example, by establishing a six-month visit interval for DEM patients and a one-year interval for MCI patients, avoiding stressing the latter ones in close visits. Our investigation represents the very first step to provide a new data-driven framework to eventually predict the diagnoses from fluency data when much of such clinical data will be available. In this regard, a Bayesian mixed effects model would be a powerful tool to get a grounded and much informative inference on the relationship between different key variables, such as the diagnosis label, the population class demographic (age, sex, and education), the semantic space (itWaC, twitter, and Wikipedia) and the value of the metrics in each semantic space. In addition, further development of this work should include a cohort of elderly healthy controls. Moreover, knowing if a patient is more performing in the density of explored concepts, many words of similar meaning, or in changing context,Max$J$, DOE, could help develop future targeted cognitive stimulation based on the value of such metrics. Cognitive stimulation [42, 43] is useful in preventing patients from abusing pharmacological therapy in favor of personalized and more targeted exercises for the maintenance of residual capacities. In other words, improving our understanding of memory retrieval task and impaired cognitive search could considerably improve the life quality of people with dementia, often prone to develop secondary diseases, such as depression [44], related to the inability to express or recall concepts. Finally, given the robustness of our results in separating healthy and nonhealthy subjects, the geometric approach could be wisely used to develop digital pretriage tools. In this way, by means of the metrics proposed in our work, patients could be divided in the two macro categories healthy and nonhealthy before the clinical examination. This would be of tremendous help in avoiding unnecessary visits to healthcare facilities. Our goal might seem ambitious and definitely challenging but maybe not so unrealistic considering the historical moment we are living in due to COVID-19. In fact, preventing most susceptible subjects to the risk of pandemic, such as those elderly people suspected of dementia, from unnecessarily going to healthcare facilities could considerably safeguard their lives.

4. Methods

4.1. Dataset. The dataset we relied on consists of semantic verbal fluencies (SVF) test records of 185 patients and 30 of healthy controls (CTR). Among them 92 patients suffer of dementia (DEM), which includes vascular dementia, frontotemporal dementia, degenerative dementia, and Alzheimer's disease, while 93 suffer of Mild Cognitive Impairment (MCI), a precursor of Alzheimer's disease. The SVF records report the sequence of Italian words, belonging to the category of animals, spoken by each patient and control subject during the test. Our work is a retrospective study of data previously collected by the Department of Mental Health, Division of Psychology, Azienda Provinciale per i Servizi Sanitari, in Trento, Italy. All the data was collected in accordance with relevant guidelines and regulations with participants' written informed consent. DEM and MCI diagnoses were made as well at Azienda Provinciale per i Servizi Sanitari of Trento, Italy, by consensus of medical specialists as geriatricians, neurologists, or psychiatrists on the basis of physiological, instrumental, and test medical data (blood tests, resonances, CT scans, liquor samples, PET, and 18 neuropsychological tests: Mini MentalState Examination, ENPA subtests, Naming, Verbal fluency on phonemic cue, Verbal fluency on semantic cue, Digit span forward/backward, Corsi span, Babcock story recall test, Rey–Osterrieth complex figure, Modified Taylor Complex Figure, Attentional matrices, Frontal assessment Battery, Clock Drawing Test, Copy of Rey–Osterrieth complex figure, Copy of Modified Taylor Complex Figure, Cornell scale for depression in dementia, Activities of daily living, and Instrumental activities of daily living). The fact of grouping together different type of dementia conditions is motivated by the small number of samples related to each dementia category we can rely on. By grouping together all dementias, we obtain a sample that is comparable with that of MCI. Anyway, for our assessment, we rely on official and specialist sources, which report: "the boundaries between different forms of dementia are indistinct and mixed forms often co-exist" (WHO, https://www.who.int/news-room/fact-sheets/detail/dementia).

The Semantic Verbal Fluency (SVF) tests were conducted at the Department of Mental Health, Division of Psychology, Azienda Provinciale per i Servizi Sanitari, in Trento, Italy, following a specific clinical protocol. In particular, the neuropsychologist asked each individual to report all words he/she can remember belonging to the category of animals, within a time interval of 60 seconds. No clues nor incentives are given to the subjects during the tests. As soon as the patient pronounces a word, the neuropsychologist takes note by hand of the spoken word, the neuropsychologist also notes the order in which the words are pronounced, and any repetitions are not marked.

4.2. Semantic Space. To define the semantic space, we leveraged on the powerful tool of word embeddings as plausible geometric proxy of such a space. Particularly, we have used three distinct word embeddings, obtained from the Italian language and generated with the popular word representation models, word2vec [27]:

(i) Italian Word Embeddings, trained on the Italian Wikipedia, and embedded in 300 dimensions [38];
(ii) itWaC, constructed from the Web limiting the crawl to the .it domain and using medium-frequency words from the Repubblica corpus and basic Italian vocabulary lists as seeds. Words are embedded in a 128-dimensional space [39];
(iii) Twitter trained on 46.935.207 tweets and embedded in a 128-dimensional space [39].

In order to get the number of animal words potentially available in the hypersphere built from the exploration radius R of spoken words for each subject, we translated into Italian the list of animals made by Greg Borenstein, available on GitHub at https://gist.github.com/atduskgreg/3cf8ef48cb0d29cf151bedad81553a54. This is used to compute the density of exploration ρ_w, as specified in the next paragraph.

4.3. Geometry. At the scale of single concepts, we provided five different indicators useful to characterize the local exploration of concepts and eventually to discriminate between healthy and nonhealthy subjects. Each subject p speaks an amount of words N during the SVF test, we call this set of words $W^p = \{w_i \text{ with } 0 \le i \le N\}$. For each patient and for the healthy controls, we define the following metrics:

(1) Maximum jump Max$J$, it is the maximum distance, in the semantic space, between two consequent words pronounced during the test. It defines the maximum instantaneous capacity to change context as follows:

$$\mathrm{Max}J^p = \max\left(\mathrm{dist}\left(w_t, w_{t+1}\right)\right)\text{where } w_t, w_{t+1} \in W^p, \tag{1}$$

where dist can be both Euclidean distance and cosine distance. Results are presented (in section Results) considering the cosine distance for all the metrics.

(2) Diameter of exploration DOE, it is the maximum distance, in the word embeddings, between the words pronounced during the test, whatever the order is; it defines the maximum capacity to change context in the whole test duration as follows:

$$\mathrm{DOE}^p = \max\Big(\mathrm{dist}\Big(w_i, w_j\Big)\Big)\text{where} \quad w_i, w_j \in W^p. \tag{2}$$

According to the measure of distance, i.e., Euclidean or cosine distance, this metric is defined, respectively, as Diameter of exploration (DOE) or Amplitude of exploration (AOE).

(3) Density of exploration ρ_w, it corresponds to the total amount of animal words potentially available in the hypersphere built from the exploration radius R, half the DOE, which has as its center the centroid C of spoken words in the semantic space as follows:

$$\rho_w^p = \sum_{a=1}^{A}\left(\mathrm{dist}\left(C^p, w_a\right) < R^p\right), \text{where } w_a \in A, \tag{3}$$

where C^p is the centroid of spoken words by patient p, for each coordinates x of the semantic space, and it is defined as follows:

$$C_x^p = \frac{1}{N}\sum_{w=1}^{N} x_{w_i}. \tag{4}$$

According to the dimension of the word embedding, the centroid will have 300 or 128 dimensions, while A is the complete set of animals in the word embeddings, and R is the radius of exploration, i.e., half the DOE.

(4) Distance d, it is the total distance covered during the test as follows:

$$d = \sum_{i=1}^{N-1} \mathrm{dist}\left(w_i, w_{i+1}\right), \text{where } w_i \in W^p. \tag{5}$$

(5) Farness far, it is the average distance of the words pronounced; it defines the ability to go far with a certain number of jumps as follows:

$$\mathrm{far} = \frac{\sum_{i=1}^{N-1}\mathrm{dist}\left(w_i, w_{i+1}\right)}{N-1}. \tag{6}$$

The significance of the above defined indicators in discriminating the three groups of subjects is evaluated by means of Kolmogorov–Smirnov statistical test and t-test, with a 95 % confidence interval (detailed results available in Tables 1 and 2 of Supplementary Materials). It is to be noticed that we are testing if each metric can separate between the three categories DEM, MCI, and healthy controls (in each word embedding). For this reason, being in a case of multiple testing, we have adjusted the p values of each performed test (Kolmogorov–Smirnov and t-test) according to the Holm–Bonferroni method.

4.4. Hierarchy. For each geometry, we provided its mesoscale organization of concepts by performing a semi-supervised clustering, using the linear algorithm of k-means, and setting the number of clusters accordingly to the elbow method (see Figure 5). Relying on these clustering configurations, we defined the explorative potential of the navigation as the total number of visited clusters and as the total number of words included in the visited clusters, for each subject. These descriptors report the total amount of clusters potentially visitable during the test and the total number of words potentially retrievable during the test. Also, for geometric indicators of previous section, even in this case, the significance of the two explorative potential metrics in discriminating the three groups of subjects has been evaluated by means of Kolmogorov–Smirnov statistical test and t-test, with a 95% confidence interval (resulting values available in Tables 2 and 3 of Supplementary Materials) and

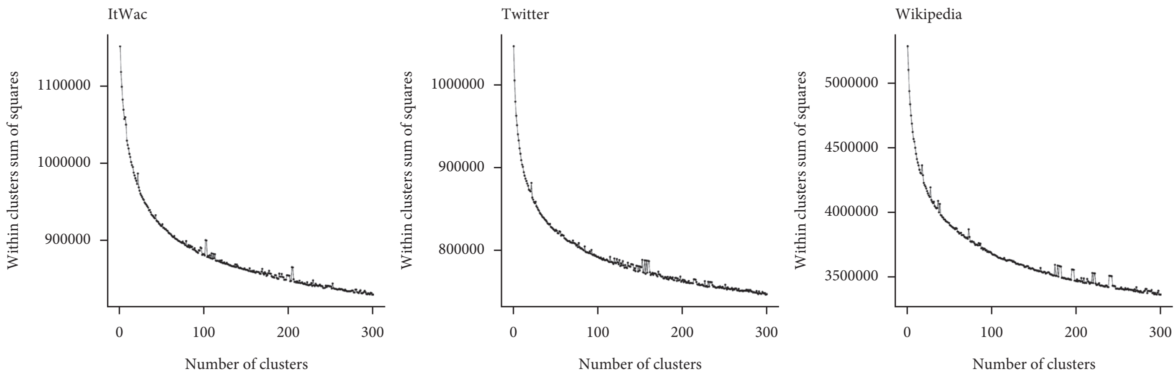

Figure 5: Elbow plots of each geometry. This figure has been generated using the publicly available R software version 3.6.3.

by adjusting p values according to the Holm–Bonferroni method. By performing a clustering analysis between the embedded visited clusters, we are able to evaluate the existence of a hierarchy in the way subjects explore concepts.

Once the clustering configuration for each geometry has been obtained, it is possible to extract the hierarchical configuration of these clusters, thanks to the coordinates of the centroids of each cluster in the geometries. In other words, the clusters are the same of k-means output, and given the position of each cluster in the geometry (identified by the centroid), it is possible to define a spatial hierarchical relationship between such clusters, in terms of distances between centroids.

Particularly, we computed the clusters' distance metric for each group setting the distance between not visited clusters equal to the double of the maximum distance between visited clusters (i.e., a proxy to infinite), in this way, we assure that the not visited clusters will not be relevant in the hierarchical analysis of that group. With the distance matrices so computed, we performed a hierarchical clustering algorithm to discover the relationship between clusters. Hierarchical relationship among clusters visited by a group is shown through dendrograms, while differences between hierarchies, i.e., in the way different groups explore concepts, are displayed through tanglegrams Figure 3. Finally, we investigated the correlation between the three groups by computing Baker's Gamma correlation coefficient in pairs for the three groups' trees (dendrograms) and by testing it against a null model. For a better understanding of the clustering analysis, we summarized below what we have done in two main steps:

(1) Clustering configuration: We defined the mesoscale organization of concepts; i.e., we identify the clusters of concepts for each category (DEM, MCI, and healthy control) and for each geometry by means of linear semisupervised clustering algorithm (k-means). These clusters represent how the expressed concepts grouped together on a semantic space and they will constitute the states of the Markov chains. Through k-means clustering, we also provided the explorative potential which defines the total amount of clusters explored by each category (see Figure 1 of Supplementary Materials).

(2) Hierarchical configuration of clusters: The clusters identified by the k-means algorithm are then given as input to a hierarchical clustering algorithm, which provides the spatial hierarchy of such clusters based on their relative distance. Intuitively, since the three categories explore different clusters, the study of the hierarchy gives us an insight into the way such clusters are explored. In order to detect any possible difference in the hierarchy, we compute the values of Baker's gamma correlation, a measure of similarity between two trees (dendrogram) of hierarchical clustering (see Figure 3).

4.5. Networks. At the macroscale, i.e., the scale of clusters of concepts, the navigation of concepts is tested by means of its Markov representation, to probe the possible alterations of mental pathways emerging from the exploration of concepts in patients with dementia. Mathematically, this corresponds to defining the transition probability from one state (i.e., cluster) to another, regardless of previously visited states. Operatively, we build three Markov chains, one for each group g, i.e. the two diagnosis and the healthy control, considering all the clusters visited by each group as the states of the Markov chain of that group, and setting the transition probabilities m equal to the relative transition frequencies from one state (r) to another (s) in each group $g$as follows:

$$m^g_{r,s} = \frac{\sum_p^{P^g} E_{(r \longrightarrow s)^p}}{\sum_p^{P^g} \sum_s^S E_{(r \longrightarrow s)^p}}, \quad (7)$$

where P^g is the total number of subjects of group g, S is the total number of visited clusters by group g, and $E_{(r \longrightarrow s)^p}$ is the outgoing edge from cluster r to cluster s for patient p^{th}. After calculating the entries $m^g_{r,s}$ we obtain, as result, the transition probability matrix $\widehat{M}$ for each category of subjects. Here, the assumption is that each subject of each category is considered as the "typical subject of that category" and corresponds to a possible realization of the typical exploration of that category. To clarify with an example, we considered 93 patients suffering of MCI; this means that the typical subject belonging to MCI has performed the tests 93 times. For practical reasons, we have transformed each transition matrix $\widehat{M}$ according to the PageRank algorithm;

this means that the stochastic process we are assuming to model the exploration of concepts behaves 85% of time according to the probabilities of the above determined Markov chain and 15% of the time according to a discrete uniform distribution [45–48] (for more details on the choice of teleportation parameter in the PageRank algorithm, we refer the reader to the dedicated section in Supplementary Materials) as follows:

$$\widehat{T} = \alpha \widehat{M} + (1-\alpha)\frac{1}{S}, \tag{8}$$

where $\widehat{T}$ represents the new transition matrix, α is equal to 0.85 according to the PageRank algorithm, and S is the number of the states of the Markov chain. Each Markov chain is then characterized by the steady state distribution $\vec{\pi}$ and by the mean first passage time matrix MFPT. Through the former, we gain information about the process at the equilibrium, while, through the latter, we can have an insight into the dynamic of the process during the exploration of concepts. Bearing in mind the memoryless property of Markov chains and that the probability of being in state r after n steps is the r^{th} entry of

$$\vec{\pi_n} = \vec{\pi_0} T^n, \tag{9}$$

where $\vec{\pi_0}$ is the probability distribution of the initial state, the steady state corresponds to the long-run equilibrium, whatever the starting state is, as follows :

$$\vec{\pi_s} = \lim_{n \longrightarrow +\infty} T^n_{r,s}. \tag{10}$$

The steady-state distribution is found by solving the system of equations obtained by imposing

$$\vec{\pi} T = \vec{\pi}. \tag{11}$$

with the constraint that all the components of $\vec{\pi}$ must sum up to 1. The steady state distribution can be obtained also by means of eigenvectors. In this case, $\vec{\pi} T = \vec{\pi}$ can be seen as

$$\vec{v} A = \lambda \vec{v}. \tag{12}$$

Therefore, $\vec{\pi}$ can be obtained from to the left-eigenvector of the square matrix T corresponding to the eigenvalue $\lambda = 1$. The MFPT is obtained from the fundamental matrix Z:

$$Z = (I - T + W)^{-1}, \tag{13}$$

where I is the identity matrix, and W is a matrix of rows identical to $\vec{\pi}$. The MFPT is determined by

$$\text{mfpt}_{i,j} = \frac{z_{j,j} - z_{i,j}}{\pi_j}. \tag{14}$$

We compare descriptors (i.e., $\vec{\pi}$ and MFPT) according to four different metrics: Pearson correlation, Spearman's correlation, covariance, and Euclidean norm of the difference, and for MFPT, we compare also its Frobenius norm. It is to be noticed that, for the MFPT, we consider the common visited states by the groups of which we want to compute the metrics and take the matrix as a vector (resulting values available in Tables 5–11 of Supplementary Materials).

Data Availability

The data that support the findings of this study are available from Azienda Provinciale per i Servizi Sanitari (APSS), Trento, Italy, but restrictions apply to the availability of these data, which were used under license for the current study and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission of APSS.

Conflicts of Interest

The authors declare no competing interests.

Authors' Contributions

Elena Bravi, Monica Dallabona, Manlio De Domenico, and Stefano Merler designed the study. Barbara Benigni prepared all the figures, performed the numerical experiments, and analyzed the data. Barbara Benigni and Manlio De Domenico wrote the main manuscript text. All authors reviewed the manuscript.

Acknowledgments

The authors thank Marco Guerini for useful discussion on word embeddings technique.

Supplementary Materials

A supplementary material file is provided with the extended results of the study and an explanation regarding the choice of teleportation parameter in the PageRank algorithm. (*Supplementary Materials*)

References

[1] M. R. Quillan, "Semantic memory," Technical Report, Bolt Beranek and Newman Inc., Cambridge, MA, USA, 1966.

[2] T. T. Hills, P. M. Todd, D. Lazer, A. D. Redish, and I. D. Couzin, "Exploration versus exploitation in space, mind, and society," *Trends in Cognitive Sciences*, vol. 19, no. 1, pp. 46–54, 2015.

[3] T. T. Hills, M. N. Jones, and P. M. Todd, "Optimal foraging in semantic memory," *Psychological Review*, vol. 119, no. 2, pp. 431–440, 2012.

[4] T. T. Hills, P. M. Todd, and M. N. Jones, "Foraging in semantic fields: how we search through memory," *Topics in Cognitive Science*, vol. 7, no. 3, pp. 513–534, 2015.

[5] E. L. Charnov, "Optimal foraging, the marginal value theorem," *Theoretical Population Biology*, vol. 9, no. 2, pp. 129–136, 1976.

[6] A. K. Troyer, M. Moscovitch, and G. Winocur, "Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults," *Neuropsychology*, vol. 11, no. 1, pp. 138–146, 1997.

[7] V. Taler and N. A. Phillips, "Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review," *Journal of Clinical and Experimental Neuropsychology*, vol. 30, no. 5, pp. 501–556, 2008.

[8] N. Linz, J. Tröger, J. Alexandersson, and A. König, "Using neural word embeddings in the analysis of the clinical

semantic verbal fluency task," in *Proceedings of the IWCS 2017—12th International Conference on Computational Semantics—Short Papers*, pp. 1–7, Montpellier, France, September 2017.

[9] J. T. Abbott, J. L. Austerweil, and T. L. Griffiths, "Human memory search as a random walk in a semantic network," *News in Physiological Sciences*, vol. 2, pp. 3050–3058, 2012.

[10] J. T. Abbott, J. L. Austerweil, and T. L. Griffiths, "Random walks on semantic networks can resemble optimal foraging," *Psychological Review*, vol. 122, no. 3, pp. 558–569, 2015.

[11] J. Borge-Holthoefer and A. Arenas, "Categorizing words through semantic memory navigation," *The European Physical Journal B*, vol. 74, no. 2, pp. 265–270, 2010.

[12] A. M. Collins and M. R. Quillian, "Retrieval time from semantic memory," *Journal of Verbal Learning and Verbal Behavior*, vol. 8, no. 2, pp. 240–247, 1969.

[13] M. Steyvers and J. B. Tenenbaum, "The large-scale structure of semantic networks: statistical analyses and a model of semantic growth," *Cognitive Science*, vol. 29, no. 1, pp. 41–78, 2005.

[14] A. Baronchelli, R. Ferrer-i-Cancho, R. Pastor-Satorras, N. Chater, and M. H. Christiansen, "Networks in cognitive science," *Trends in Cognitive Sciences*, vol. 17, no. 7, pp. 348–360, 2013.

[15] C. S. Siew, D. U. Wulff, N. M. Beckage, and Y. N. Kenett, "Cognitive network science: a review of research on cognition through the lens of network representations, processes, and dynamics," *Complexity*, vol. 2019, Article ID 2108423, 24 pages, 2019.

[16] M. N. Jones, T. T. Hills, and P. M. Todd, "Hidden processes in structural representations: a reply to Abbott, Austerweil, and Griffiths," *Psychological Review*, vol. 122, no. 3, pp. 570–574, 2015.

[17] J. C. Zemla, Y. N. Kenett, K.-S. Jun, and J. L. Austerweil, "U-invite: estimating individual semantic networks from fluency data," *CogSci*, vol. 1, no. 1, pp. 36–58, 2016.

[18] J. C. Zemla and J. L. Austerweil, "Estimating semantic networks of groups and individuals from fluency data," *Computational Brain and Behavior*, vol. 1, no. 1, pp. 36–58, 2018.

[19] N. B. Mota, N. A Vasconcelos, N Lemos et al., "Speech graphs provide a quantitative measure of thought disorder in psychosis," *PLoS One*, vol. 7, Article ID e34928, 2012.

[20] A. S. Chan, N. Butters, J. S. Paulsen, D. P. Salmon, M. R. Swenson, and L. T. Maloney, "An assessment of the semantic network in patients with Alzheimer's disease," *Journal of Cognitive Neuroscience*, vol. 5, no. 2, pp. 254–261, 1993.

[21] A. J. Lerner, P. K. Ogrocki, and P. J. Thomas, "Network graph analysis of category fluency testing," *Cognitive and Behavioral Neurology*, vol. 22, no. 1, pp. 45–52, 2009.

[22] L. Bertola, N. l. B. Mota, M. Copelli et al., "Graph analysis of verbal fluency test discriminate between patients with Alzheimer's disease, mild cognitive impairment and normal elderly controls," *Frontiers in Aging Neuroscience*, vol. 6, p. 185, 2014.

[23] M. Stella and Y. N. Kenett, "Viability in multiplex lexical networks and machine learning characterizes human creativity," *Big Data and Cognitive Computing*, vol. 3, no. 3, p. 45, 2019.

[24] A. P. Christensen, Y. N. Kenett, K. N. Cotter, R. E. Beaty, and P. J. Silvia, "Remotely close associations: openness to experience and semantic memory structure," *European Journal of Personality*, vol. 32, no. 4, pp. 480–492, 2018.

[25] F. Günther, L. Rinaldi, and M. Marelli, "Vector-space models of semantic representation from a cognitive perspective: a discussion of common misconceptions," *Perspectives on Psychological Science*, vol. 14, no. 6, pp. 1006–1033, 2019.

[26] P. Mandera, E. Keuleers, and M. Brysbaert, "Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: a review and empirical validation," *Journal of Memory and Language*, vol. 92, pp. 57–78, 2017.

[27] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Advances in Neural Information Processing Systems*, vol. 26, pp. 3111–3119, 2013.

[28] A. I. Tröster, D. P. Salmon, D. McCullough, and N. Butters, "A comparison of the category fluency deficits associated with Alzheimer's and Huntington's disease," *Brain and Language*, vol. 37, no. 3, pp. 500–513, 1989.

[29] J. D. Henry, J. R. Crawford, and L. H. Phillips, "Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis," *Neuropsychologia*, vol. 42, no. 9, pp. 1212–1222, 2004.

[30] P. Garrard, M. Lambonralph, K. Patterson, K. Pratt, and J. Hodges, "Semantic feature knowledge and picture naming in dementia of Alzheimer?s type: a new approach," *Brain and Language*, vol. 93, no. 1, pp. 79–94, 2005.

[31] D. P. Salmon and M. W. Bondi, "Neuropsychological assessment of dementia," *Annual Review of Psychology*, vol. 60, no. 1, pp. 257–282, 2009.

[32] Z. Shao, E. Janse, K. Visser, and A. S. Meyer, "What do verbal fluency tasks measure? predictors of verbal fluency performance in older adults," *Frontiers in Psychology*, vol. 5, p. 772, 2014.

[33] J. Borge-Holthoefer, Y. Moreno, and A. Arenas, "Modeling abnormal priming in Alzheimer's patients with a free association network," *PLoS One*, vol. 6, 2011.

[34] A. G. Huth, W. A. De Heer, T. L. Griffiths, F. E. Theunissen, and J. L. Gallant, "Natural speech reveals the semantic maps that tile human cerebral cortex," *Nature*, vol. 532, no. 7600, pp. 453–458, 2016.

[35] A. Puglisi, A. Baronchelli, and V. Loreto, "Cultural route to the emergence of linguistic categories," *Proceedings of the National Academy of Sciences*, vol. 105, no. 23, pp. 7936–7940, 2008.

[36] L. Manubens-Gil, L. Manubens-Gil, S. Jim et al., "Neuromorphological and wiring pattern alterations effects on brain function: a mixed experimental and computational approach," *Frontiers in Systems Neuroscience*, vol. 10, 2015.

[37] T. L. Griffiths, M. Steyvers, and J. B. Tenenbaum, "Topics in semantic representation," *Psychological Review*, vol. 114, no. 2, pp. 211–244, 2007.

[38] Retrieved on February 2020 from Italian Word Embeddings, https://hlt.isti.cnr.it/wordembeddings/.

[39] A. Cimino, L. De Mattei, and F. Dell'Orletta, "Multi-task learning in deep neural networks at EVALITA 2018," in *Proceedings of the 6th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA'18)*, pp. 86–95, Accademia University Press, Naples, Italy, December 2018.

[40] J. Goñi, I. Martincorena, B. Corominas-murtra, G. Arrondo, S. Ardanza-trevijano, and P. Villoslada, "Switcher-random-walks: a cognitive-inspired mechanism for network exploration," *International Journal of Bifurcation and Chaos*, vol. 20, no. 3, pp. 913–922, 2010.

[41] F. B. Baker, “Stability of two hierarchical grouping techniques case i: sensitivity to data errors,” *Journal of the American Statistical Association*, vol. 69, no. 346, pp. 440–445, 1974.

[42] V. Orgeta, P. Leung, L. Yates et al., “Individual cognitive stimulation therapy for dementia: a clinical effectiveness and cost-effectiveness pragmatic, multicentre, randomised controlled trial,” *Health Technology Assessment*, vol. 19, no. 64, pp. 1–108, 2015.

[43] A. Bahar-Fuchs, A. Martyr, A. M. Goh, J. Sabates, and L. Clare, “Cognitive training for people with mild to moderate dementia,” *Cochrane Database of Systematic Reviews*, vol. 2, no. 1, Article ID 30909318, 2019.

[44] B. W. Rovner, J. Broadhead, M. Spencer, K. Carson, and M. F. Folstein, “Depression and Alzheimer’s disease,” *American Journal of Psychiatry*, vol. 146, pp. 350–353, 1989.

[45] T. L. Griffiths, M. Steyvers, and A. Firl, “Google and the mind,” *Psychological Science*, vol. 18, no. 12, pp. 1069–1076, 2007.

[46] L. Page, S. Brin, R. Motwani, and T. Winograd, “The Pagerank citation ranking: bringing order to the web,” Technical Report, Stanford InfoLab, Stanford, CA, USA, 1999.

[47] R. Mihalcea, P. Tarau, and E. Figa, “Pagerank on semantic networks, with application to word sense disambiguation,” in *2004 Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*, pp. 1126–1132, ACM Library, New York, NY, USA, August 2004.

[48] S. A. Sloman, B. C. Love, and W.-K. Ahn, “Feature centrality and conceptual coherence,” *Cognitive Science*, vol. 22, no. 2, pp. 189–228, 1998.