Cognitive networks for knowledge modelling: A gentle tutorial for data- and cognitive scientists

*Edith Haim and Massimo Stella**

CogNosco Lab, Department of Psychology and Cognitive Science, University of Trento, Italy

**Corresponding author:* massimo.stella-1@unitn.it

Abstract

In this tutorial paper, we discuss cognitive networks as powerful models for understanding human cognition and knowledge. Cognitive networks are representations of associative knowledge between concepts in a cognitive system apt at acquiring, storing, processing and producing language, i.e. the mental lexicon. In a cognitive network, nodes represent concepts with links expressing relations, such as semantic, syntactic, phonological and visual connections, e.g. "canine" and "dog" (nodes) linked by "being synonyms" (link). Hence, cognitive networks represent associative knowledge in mathematical, measurable and quantifiable ways. Can such structure be used to gain insights over cognitive phenomena? We explore this research question by reviewing recent, pioneering key applications and limitations of cognitive networks across visual, auditory, and semantic language processing tasks, either in healthy or clinical populations. We also review applications of cognitive networks modelling language acquisition, reconstructing text content and assessing creativity or personality traits in individuals. Our tutorial also gently introduces the reader to mathematical notations, definitions and measures about single-layer and multiplex networks as well as hypergraphs. Last but not least, across phonological, semantic and syntactic networks, we guide the reader through relevant psychological frameworks, datasets and software packages that might all aid current and future cognitive network scientists.

1. Introduction

The cognitive reflection of language in the mind is a complex system where several idealised units interact together at multiple scales to convey knowledge, meaning and emotions (Tulving 1972; Aitchison 2012; Zock 2020; Hills & Kenett 2021). The complexity of such a cognitive system originates from the fact that "more is different": Combining ideas with specific features can create novel cognitive entities with distinctive characteristics, e.g. "clause" (proposition) with "santa" (saint). Complexity is not limited to the spheres of meaning or emotions but rather extends also to other aspects of knowledge. For instance, we have the ability to remember and piece together phonemes (i.e. sound units) to form words as phoneme sequences (Vitevitch 2008). We can also stack words to create sentences. When combined, multiple sentences create narratives that shape ideas, viewpoints, emotions, and contexts, making all these aspects highly interconnected. All this information is stored and processed across several cognitive mechanisms, either sequentially or in parallel (Aitchison 2021; Doczi 2019). These idealised linguistic forms are structured representations of knowledge as encoded within human cognition. Other examples include mental images or conceptual frameworks, e.g. visualising a red apple when thinking about the concept "apple" (Ciaglia et al. 2023). Notice that the terms "words" and "concepts" are distinct - words represent linguistic units whereas concepts are cognitive reflections of ideas in the human mind (Vigliocco et al. 2018). However, despite this distinction, since our focus is on modelling cognition via models yet unable to distinguish between concepts and words, we use these two terms interchangeably.

If concepts can be assembled together, shaping our knowledge systems, the necessity arises for models that can elucidate, quantify, and explore such assembling or associative structures (Doczi 2019; Hills & Kenett 2021). These models should also focus on highlighting the interplay between associative structure and other cognitive phenomena, dynamics and functions, grounded on such organisation of concepts and ideas (Stella et al. 2017). In pursuit of this goal, cognitive modelling introduced the metaphor of a mental lexicon (Collins & Loftus 1975; Aitchison 2012). Contrary to its name, this is not a mere dictionary but rather a complex cognitive system, comprising multiple interacting elements. The mental lexicon mostly includes semantic memory, acting as a repository for linguistic knowledge about concept meanings (Sizemore et al. 2018; Kenett et al. 2018; Kumar 2021), alongside other subsystems handling phonological knowledge (Vitevitch 2008), visual cues (Kennington & Schlangen 2015), and more (Doczi & Kormos 2015; Zock 2020). In this lexicon, mental representations can share similar features, forming a "cobweb" of interactions influencing knowledge acquisition, storage, processing, and production (Steyvers & Tenenbaum 2005; Aitchison 2012). This analogy aligns with the concept of a complex network (Collins & Loftus 1975) but with a unique twist: links in the mental lexicon are "invisible" unlike those in physical systems, making the structure of the mental lexicon inaccessible for direct, exact reproduction in a lab setting. For instance, whether a researcher might explore and tinker with brain connections in the anatomy of a real brain in a given lab setting, researchers cannot directly observe whether two ideas are encoded as synonyms in an individual's mental lexicon but rather need to rely on external tasks or definitions, i.e. indirect probing.

Consequently, the structure of the mental lexicon has to be indirectly investigated through cognitive tasks. A key issue with these tasks is that they represent linguistic data without providing insights into the cognitive structure of associative knowledge behind the data itself. For instance, the data coming from a fluency task (see Section 2.6) remains a sequence of words, with no explicit information about relationships between them. Hence, accessing only data from cognitive tasks still misses information about how concepts are organised within human cognition. This research gap is filled by network models, which act as representation proxies of the structural organisations of concepts in the mental lexicon, with all due limitations relative to using proxies (Kenett & Hills 2022).

Notice that cognitive networks, the main topic of this tutorial, are not artificial neural networks (Fatima et al. 2021). The latter are interconnected computational units, i.e. artificial neurons, which can integrate or modify signals over time, thus possessing a distinctive computational power, e.g. reproducing the OR logic. The cognitive networks reviewed in this tutorial paper are rather representational models, i.e. they explicate the layout of conceptual relationships in associative knowledge, even in absence of computational elements. Despite this distinction, cognitive network science has grown over the years as a data-centric field, employing multidisciplinary techniques from cognitive science, psychology, social science, mathematics, physics, statistics, and computer science (Hills et al. 2009; Kenett et al. 2018; Siew et al. 2019; Siew & Castro 2020; Stella 2022).

This tutorial paper integrates excellent pre-existing reviews (Siew et al. 2019; Siew & Castro 2020) by focusing on: (i) outlining the most recent advancements and cognitive interpretations in the field focusing on network science methods and relevant literature as to appeal to computational scientists, and (ii) reviewing the latest large-scale datasets and coding packages that can support computational scientists in using cognitive networks as models for understanding human cognition and behaviour.

2. Basic Definitions

Let us start by defining in simple terms what is a cognitive network:

Definition 1: *A cognitive network is a data-informed model explicating associations (links) between cognitive representations of concepts (nodes).*

Being data-informed implies that cognitive networks draw their insights and structure from empirical data and psychological models (Siew et al. 2019; Stella 2022), reflecting real-world patterns and interactions in combination with underlying theoretical assumptions and psychological constructs. The phrase “explicating associations” underscores the networks´ ability to account for one or several connections or relationships present between concepts in the mental lexicon, e.g. phonological similarities (Vitevitch 2008), memory recall patterns (De Deyne et al. 2013), syntactic dependencies (Stella 2020) and so on. Furthermore, “cognitive representations of concepts” pertains to the mental constructs that individuals hold for various ideas or entities, ranging from meanings to visual, phonological, emotional, and syntactic features, among others (Doczi & Kormos 2015).

Examples of nodes in cognitive networks include various linguistic elements, such as semantic representations of words or idioms (one or more words conveying a unique meaning, e.g. “leap of faith”). Morphemes capturing the root or stem of a word encode another type of node (e.g. “happi” being the stem of “happiness”), as well as phonological transcriptions of words (see Fig. 1) and many others (Doczi & Kormos 2015; Sizemore et al. 2018; Valba et al. 2021). Conceptual associations within cognitive networks encompass a diverse array of aspects within the mental lexicon and include semantic, phonological, syntactic, sensorimotor, affective, and visual aspects of the mental lexicon, among many others (Martinčić-Ipšić et al. 2016; Stella et al. 2018; Siew et al. 2019; Castro & Siew 2020; Levy et al. 2021; Stella 2022). For instance, semantic associations involve connecting words like “apple” with concepts sharing some features or overlapping in meaning, like “fruit” or “red” (Doczi 2019). Phonological associations might be observed in recognising sound similarities between words such as “cat” and “cab” (Vitevitch 2008). Syntactic associations involve understanding grammatical structures, as demonstrated in sentences like “The cat looks at the cab” (Stella et al. 2022). Moreover, affective associations involve linking emotions to words, such as associating positivity with “celebration” or negativity with “failure” (Semeraro et al. 2022). Lastly, mental imagery can be evoked through visual associations - like imagining a beach, palm trees, and waves when encountering the word “vacation” - and sensorimotor associations, illustrated by mentally envisioning actions like riding a bicycle when the word “bike” is mentioned (Ciaglia et al. 2023).

Even psychometric networks (Golino & Epskamp 2017) would fall into Definition 1, although their link construction is considerably more advanced (Golino et al. 2020) than the scope of this tutorial paper. In fact, psychometric networks encapsulate correlations between numerical sets (e.g. responses to items) and are thus not equivalent to cognitive networks of conceptual associations.

Some pioneering approaches (Stanghellini et al. 2023) are entwining the structures of cognitive and psychometric networks, merging the semantic content of items in psychometric questionnaires with numerical responses. Notice that Definition 1 does not include brain networks (Beaty et al. 2018), which can be used to predict cognitive phenomena or traits (e.g. creativity levels) but encode relationships between brain circuits rather than between cognitive representations of concepts.

2.1. Simple mathematical definitions: Edges, vertices and sets

When considering only a single type of conceptual association (Siew et al. 2019; Stella 2022), cognitive networks can be represented as a couple of sets (V, E), which includes a vertex set $V = \{i, \ldots, s, \ldots, j\}$ and an edge set E, e.g. $\{(i, j), \ldots (s, i)\}$. In undirected conceptual associations, the relationship between two words, represented as the link (i, j), is symmetrical, meaning it is equivalent to the reverse order (j, i). This symmetry is denoted as i-j, indicating that the connection between the two words is bidirectional. For example, considering the synonyms "land" and "earth", the undirected link "land"-"earth" conveys that both terms are conceptually related with no directionality or hierarchy between them. Importantly, in undirected links, the order in which words are connected is irrelevant, emphasising the mutual and interchangeable nature of these types of associations.

Instead, in directed associations the order of words is relevant (Stella et al. 2018). For instance, in the sentence "birds are a broader category of doves", the fact that "bird" is a hypernym of "dove" could be coded as a directed connection $bird \rightarrow dove$. Specifically, this coding signifies that "bird" is a broader category encompassing the more specific term "dove". The directional arrow (→) captures the hierarchical flow, from the hypernym "bird" to the hyponym "dove". This emphasises the directional relationship between these two words. Such a connection could be represented as a couple of elements in case of considering the order, i.e. the source node will appear first in the list, followed by the target node. Hence, $bird \rightarrow dove$ could be represented as $(bird,\ dove)$. Instead, $(dove,\ bird)$ or, equivalently, $dove \rightarrow bird$, would not encode a hypernym but rather a hyponym.

Networks where all links are undirected are called undirected networks (Newman 2018). Networks with only directed links are called directed networks. Networks combining directed and undirected connections can be recast as special cases of directed networks in case undirected connections are split in reciprocated directed connections, for instance the relationship "good"-"positive" could be split in "good"→"positive" and the reciprocating connection "positive"→"good". Figure 1 reports an undirected cognitive network, with a vertex set

$V = \{cat, can, cab, man, crab\}$. The links between the entries map phonological similarities between their transcriptions. Such phonological similarities regard words differing by the addition, substitutions or deletion of one phoneme (Vitevitch 2008). For instance, considering the vertex set of Figure 1, “cab” can be changed into “crab” by adding an /r/ sound, while substituting the /c/ in “can” for /m/ results in “man”. These examples showcase how slight phonological modifications can result in a series of words with interconnected phonological similarities, i.e. sounding similar with each other.

In mathematical set notation (Newman 2018), the set of phonological links in Figure 1 would be represented as $E = \{(cat, cab), (can, cab), (can, cat), (man, can), (cab, crab)\}$, indicating the relationships among the words in terms of phonological similarities (Turnbull 2021). What is reported in Figure 1 is also called a “single-layer cognitive network” because it encapsulates links of the same nature (Stella et al. 2017).

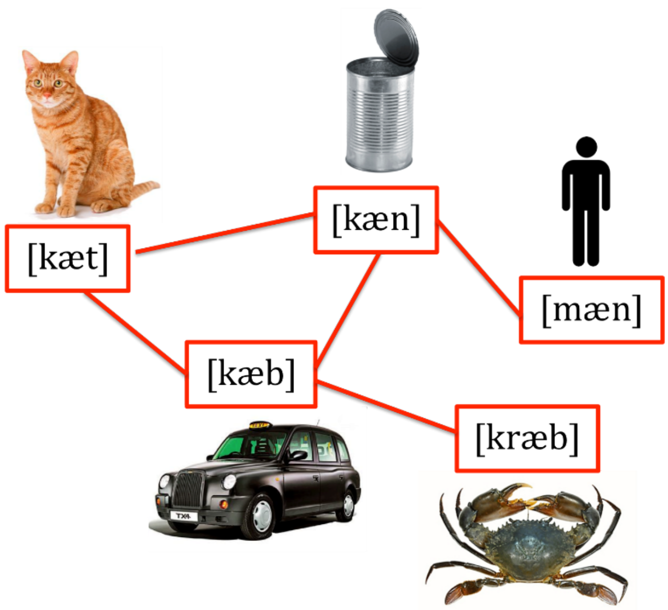

Figure 1: A cognitive network where nodes represent phonological transcriptions of English words in the IPA alphabet (see also Turnbull 2021). The links between transcriptions indicate phonological similarities (i.e. two words differing for the addition, substitution or deletion of one phoneme only and thus sounding similar to each other).

2.2. Single-layer and multiplex lexical networks

Extensive psycholinguistic research has shown that among the very same set of concepts there can be multiple types of conceptual associations (Fay & and Cutler 1977; Aitchison 2012; Abrams

& Davis 2016). In other words, the concepts can relate to one another in multiple ways. For instance, there could be semantic similarities between "cat" and "crab" in terms of both being animals or syntactic relationships between "cat" and "man" as encoded in a given sentence, e.g. "The man feeds the cat." This existence of multiple types of conceptual associations among the same set of concepts leads to edge-coloured graphs known as *multiplex lexical networks* (Stella et al. 2017; Stella et al. 2018; Stella 2020b). In multiplex lexical networks diverse layers capture various dimensions of connectivity among the same set of concepts with relationships including semantic, syntactic, phonological links and many other types. Figure 2 reports an example with only two layers. In general, in a multiplex lexical network, multiple edge-lists $E_1, ..., E_L$ co-exist, each one composing what is called a "network layer". Historically, the idea of different layers originated in social science contexts for mapping different types of social relationships among the same set of individuals (Newman 2018). In cognitive science, instead, relationships of different colours indicate different aspects of associative knowledge being simultaneously or sequentially used for acquiring, storing, and processing concepts (Hills et al. 2009; Citraro et al. 2023). For instance, the phenomenon of malapropism (Fay & Cutler 1977) occurs when someone can access the phonological information relative to a given target word but then fails at activating the correct semantic information that would lead to using the target word in the correct context. An example of malapropism could be "I band but I don´t break" where we use "band" instead of "bend". The occurrence of malapropism in common language signifies that phonological and semantic layers are distinct but interact with each other during language processing and production (see also Doczi 2019). This importance is further underlined by another phenomenon rising from the interactions between semantic and phonological aspects of the mental lexicon, i.e. the so-called "tip of the tongue" event (Abrams & Davis 2016). This takes place when someone can access the semantic features of a given target word but is unable to retrieve and produce the phonological instruction for naming the target. For instance, someone might be familiar with the concept of a crane (e.g. being a bird, having wings, flying) but ultimately being unable to say the name at the moment and rather use a common English expression: "I have it on the tip of my tongue." Malapropisms and tip-of-the-tongue events both underline how important it is to identify models that can encapsulate at the same time semantic and phonological relationships between concepts, like multiplex lexical networks.

Multiplex networks where there are no explicit inter-layer links are also called "edge-coloured" graphs. This naming was motivated by links being drawn in different colours, indicating different types of relations between nodes (Stella et al. 2017; De Domenico 2022). For example, Figure 2 portrays an edge-coloured graph mixing phonological associations (highlighted in red) and semantic associations (highlighted in cyan). Edge-coloured graphs and multiplex lexical networks (Stella et al. 2018) can be used as synonyms whenever there are no explicit costs for transitioning between layers. In cognitive networks it is unclear how to model interactive or subsequent transitions between phonology and semantics so that most multiplex lexical networks currently present in the literature are edge-coloured graphs.

Why use multiplex rather than simpler single-layer networks? The combination of network layers might highlight phenomena that could not be observed in individual networks. For instance, in (Siew & Vitevitch 2019) orthographic and phonological similarities highlighted facilitative effects in visual word recognition that were not observed in orthographic similarities or phonological similarities separately. Other examples include phenomena such as enhanced preferential acquisition (Stella et al. 2017) and improved lexical processing based on word distance in clinical populations (Castro & Stella 2019; Castro et al. 2020; Baker et al. 2023). These phenomena emerged only when semantic and phonological layers were combined in the multiplex structure: Multiplexity might open up novel perspectives when modelling semantic memory. For a more comprehensive review on the topic, we refer to (Stella et al. 2022).

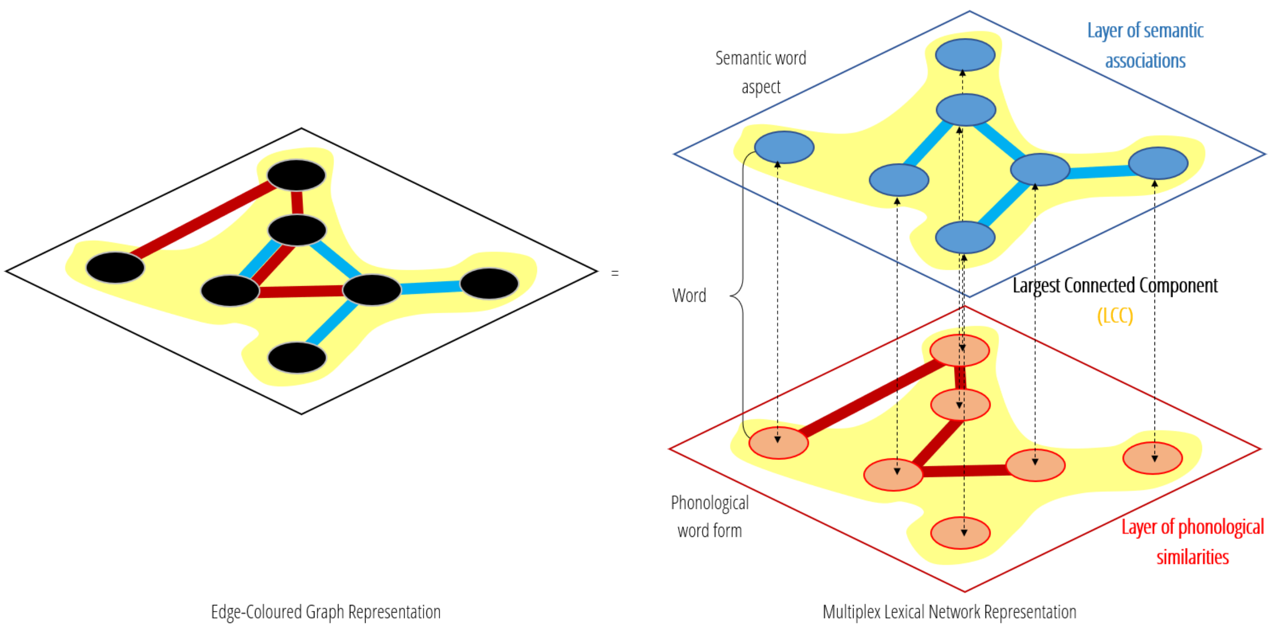

Figure 2: A multiplex lexical network where nodes engage in phonological (red) and semantic (cyan) associations. These networks can be visualised by either emphasising their multi-layer structure (right, multiplex representation), where all nodes are replicated across layers, or by collapsing layers together while preserving different colours (left, edge-coloured representation). The largest connected component of the network is highlighted in yellow.

2.3. Simple mathematical definitions: Adjacency matrices

Both single- and multiplex networks can be represented as matrices (Newman 2018; De Domenico 2022). Figure 3 depicts an example of a matrix for a single-layer network. An adjacency matrix S encodes the connectivity of any single-layer network in its elements, e.g. s_{ij} where each element represents the presence (1) or absence (0) of a connection between nodes i and j. If there is a link between i and j then $s_{ij} = 1$, otherwise $s_{ij} = 0$ (see Fig. 3). This also means that the dimensionality of S is $N \times N$, where $N = |V|$ is the number of nodes in the network. In undirected networks, S must be symmetric, i.e. $s_{ij} = s_{ji}\ \forall i, j \in V$. S encapsulates the so-called "topology" of a network, i.e. how links are organised between nodes. Elements s_{ii}, when present, are called self-loops. In presence of weights, W contains elements w_{ij} set as the weight for (i, j), e.g. the number of times i and j co-occurred in text. An undirected, unweighted network without self-loops is also called a "simple network" (Newman 2018).

	A	**B**	**C**	**D**
A	0	1	1	0
B	1	0	1	0
C	1	1	0	1
D	0	0	1	0

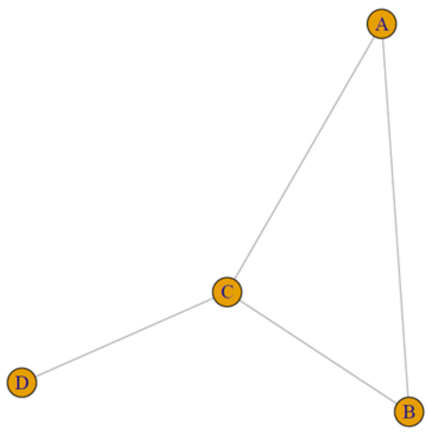

Figure 3: Example of an undirected single-layer network (right) and its adjacency matrix (left). The existence or absence of connections between the entries A, B, C and D is visualised as either 1 (link exists) or 0 (no link exists).

2.4. Simple mathematical definitions: Degree, power-laws, hubs and degree distributions

In simple networks, summing over the i-th column or row of matrix S provides the total number of links featuring node i. This total is known as the network degree k_i. The network degree k_i is a local network measure specific to node i: Degree counts only interactions that directly consider one node while neglecting the rest. In phonological networks, network degree has long been called phonological neighbourhood density (PND), counting words which are phonologically similar to a given target (Vitevitch 2008; Turnbull 2021). Degree/PND has been shown to influence lexical identification in confusability tasks and even short-term retainment (cf. Vitevitch & Mullin 2021). For example, consider a node in a semantic network with a high degree, meaning it is connected to numerous other nodes. In a confusability task, i.e. say whether a sequence of phonemes represents a real word, participants might find it challenging to quickly and accurately identify this node due to the abundance of interconnected associations, leading to potential confusion with closely related concepts. Conversely, in short-term retention tasks, e.g. memorise and reproduce words, a node/word with a high degree may be recalled more easily, as its rich connectivity aids in creating a robust and interconnected mental representation, facilitating easier retrieval from memory according to a spreading activation phenomenon (Vitevitch & Mullin 2021).

Much attention has been devoted to characterising a network based on the statistical distribution of its degrees (see also Steyvers & Tenenbaum 2005). Consider the probability $p(K = k)$ of sampling one node in a network and finding it has a degree equal to k. Sampling means considering one node at random by picking it among all available nodes in a given network structure. Thus, $p(K = k)$ is the fraction of nodes with that degree. If the probability distribution of finding a node with a specific degree follows a power-law distribution, then the network is said to be a power-law one. This power-law distribution can be noted as $p(K = k) = z \cdot k^{-\alpha}$ where z is a normalisation factor and α a power-law exponent. In cognitive networks, a power-law distribution might manifest as the majority of concepts having a low degree of semantic connections, while a few concepts gather a vast amount of connections or associations. Fitting a power-law distribution directly to $p(K = k)$ suffers from high error margins over nodes with infrequently high degrees (cf. Alstott et al. 2014). This issue can be reduced if one considers the cumulative degree distribution $C(K \geq k)$, i.e. the probability of finding a node with a degree equal to or higher than k. Notice, however, that fitting a specific equation to an empirical degree distribution can be daunting not only in statistical terms but also for the interpretation of degree distributions. Power-law degree distributions are interesting mostly because they indicate the presence of a few hub nodes, i.e. nodes with high degrees, thanks to their heavy-tailed decay. The “tails” of a distribution refer to the extreme ends of the distribution, where the probability of observing values becomes progressively smaller. A distribution can have a “heavy tail” if the probability of extreme events, located in a tail, decreases more slowly than it would in a normal or exponential distribution. This implies that rare and high-magnitude events, e.g. hubs, occur more frequently than expected in heavy-tailed rather than non-heavy-tailed distributions (e.g. Gaussian distributions). Hence power-law degree distributions in networks imply that a few nodes have exceptionally high degrees (Champagnat et al. 2013). However, this can be the case also in presence of other heavy-tailed distributions, e.g. log-Gaussian distributions or stretched exponentials. Since from a cognitive perspective it remains unclear whether cognitive representations have a specific benefit in displaying power-law degree distributions, the emphasis should be put not on detecting power-lawness but rather in identifying hubs. For large values of k, the probability of finding a node with a high degree, i.e. a hub, can be considerably higher for a heavy-tailed than for an exponential distribution with the same average degree $\langle k \rangle$. For example, in a semantic network with a heavy-tailed degree distribution, a small number of concepts - yet higher than in random graphs with exponential degree distributions - may serve as prominent hubs, forming numerous connections, while the majority of concepts have fewer connections (De Arruda et al. 2017; Newman 2018).

In semantic networks, hubs often represent pivotal concepts that play significant roles in information flow and cognitive processing (Stella 2020b). These hubs may correspond to central ideas or key semantic components that hold considerable influence over the overall network structure (De Arruda et al. 2017). A commonly used non-parametric definition identifies hubs as nodes falling in the upper percentiles of the degree distribution, e.g. nodes with degrees in the 95^{th} percentile of $p(K = k)$ (cf. Stella 2020b and De Domenico 2022). By identifying and understanding these hubs, cognitive network scientists can gain insights into the organisational principles of cognitive networks, discerning the key nodes that shape information flow and contribute substantially to cognitive processes. For instance, hubs might be general-level cues that mediate memory recollections of many other concepts (De Deyne et al. 2013). It was recently shown that hubs tend to be acquired earlier by normative English-speaking toddlers (Hills et al. 2009; Sizemore et al. 2018), whereas the acquisition of semantic hubs is impaired in late learners (Beckage et al. 2011). For example, a normative English-speaking toddler might readily acquire a semantic hub such as “apple” due to its common use and broad applicability in everyday contexts. In contrast, a late learner might struggle to acquire such foundational concepts, leading to difficulties in comprehending fundamental terms that serve as central nodes in their cognitive network which further interferes negatively with their ability to learn and connect other words in their semantic network (Beckage et al. 2011; Hills et al. 2009; Sizemore et al. 2018).

2.5. A bit less simple mathematical definitions: Network Laplacians and spreading activation

For any single-layer simple network, let us consider a matrix D with node degrees on its main diagonal and 0s everywhere else. Let us also consider the adjacency matrix S of this network. Through S and D we can define a network Laplacian as $L = D - S$ (cf. Koponen 2021), which encodes information about how much activation flow can spread along network links.

Let us discuss diffusive processes through a physical metaphor. Imagine a liquid with droplets flowing along links on a network structure relative to the Laplacian L. The main diagonal of L comes from D (for simple networks) and it encodes information over the total flow that can accumulate across links surrounding each node, e.g. one droplet per link. However, L contains also entries from the adjacency matrix S so each off-diagonal entry in L indicates where droplets can arrive into or leave nodes when flowing through connections. Overall, it is expected that L regulates the diffusion of droplets across links and over nodes in a given network (Newman 2018).

The Laplacian L is also relevant to cognitive scientists modelling spreading activation in lexical retrieval (Siew 2019). Spreading activation was introduced by Collins and Loftus to model semantic priming (Collins & Loftus 1975). A prime/node i receives an initial activation level $a_i(0)$. At the next time step, the activation $a_i(0)$ spreads uniformly across the k_i links of i to its neighbours. When the activation spreads, at each time step the activation level from a given node - or part of it - is transmitted to its connected neighbours, influencing their subsequent activation levels. The signal can receive dampening or spread integrally based on factors such as the strength of connections between nodes. For instance, when flowing along connections, a part of the activation signal might get lost (uniform dampening). It might also be that some conceptual associations might be weaker and thus receive lower activation units, whereas other associations might be stronger and thus able to channel stronger activation signals (Collins and Loftus considered this latter case as “highways” able to withstand higher traffic loads, see Collins & Loftus 1975). Iteratively, the activation signals of all nodes with some activation levels have to spread across neighbours. The process continues iteratively until a final time step is reached, typically when the system achieves a stable state or a desidered number of iterations have been reached (Collins & Loftus 1975; Castro & Siew 2020). Simulated activation levels on phonological networks were shown to replicate key findings in lexical processing (Vitevitch & Mullin 2021), indicating a crucial link between this model and some instances of knowledge processing in the mental lexicon.

Importantly, network structure can influence the amount of activation a target node can receive, thus influencing its cognitive processing. Clustering, degree and distance (see down below) can promote the concentration of activation over nodes related to the prime and thus suggest potential candidates for recollection (Siew 2019; Kumar et al. 2021). In cognitive networks representing the associative knowledge of students in the history of science, an appropriately normalised Laplacian was shown to capture nodes representing prominent scientists (Koponen 2021).

2.6. Generalisations of degree and Laplacians to multiplex lexical networks

The above notions can all be extended to multilayer networks. In the case of this tutorial paper, let us focus on the simpler instance of multiplex networks (Battiston et al. 2017). A super-adjacency matrix and a super-Laplacian matrix can be built as block matrices, with the diagonal elements being set to the adjacency (Laplacian) of individual layers and off-diagonal elements being equal to diagonal matrices expressing inter-layer relationships. For more details we refer the interested reader to (De Domenico 2022). Applying multilayer or multiplex Laplacians to cognitive networks remains a scarcely explored research direction.

Instead, extending the idea of degree to multiplex networks is relatively simpler. Summing degrees across layers, the multidegree m_i of node i is defined as:

$$m_i = \sum_{\alpha} k_i^{(\alpha)},$$

where $k_i^{(\alpha)}$ is the degree of node i in layer α, e.g. the number of cognitive links of colour/layer α surrounding node i (Stella et al. 2017). In multiplex networks, α can enumerate the different replicas of a node across each layer, i.e. replica nodes. In Figure 2, one word could have two replica nodes, one on the phonological layer and one on the semantic layer. In Figure 2 (see Section 2.2.), the highest multidegree is 4. A node having a multidegree of 4 means that, when considering cognitive links across different layers, that node is connected to four distinct other nodes across all layers. The multidegree of a node might thus be different from the single-layer degrees of the node's replicas. Multidegree can be used to define multiplex hubs: Nodes in the 95th percentile of the multidegree distribution. Stella (2020b) showed that a multiplex network representation of the mental lexicon with 16,000 English words over 4 semantic/phonological layers relied heavily on multidegree hubs to preserve connectivity (and thus a chance for activation spreading) among large portions of words. Removing these hubs made the network highly fragmented and disconnected, thus preventing activation spreading from reaching several network regions.

In semantic multiplex lexical networks, multidegree is also called semantic richness, i.e. the number of concepts syntactically or semantically related to a target node/concept (Aitchison 2012; Stella 2020). Nodes with higher semantic richness correspond to more general concepts, potentially being used as synonyms to several other concepts, or words occurring across multiple contexts and thus acquiring more syntactic links.

2.7. Beyond degree: Local clustering coefficient and global clustering coefficient

Whereas degree is a local feature of nodes, local clustering and closeness centrality, respectively, capture meso-scale (intermediate-sized structures or patterns within the network) and global aspects (overall structure and efficiency of information flow across the entire network) of network topology (Siew et al. 2019).

Local clustering is a meso-scale measure in that it characterises a node based on links between its neighbours. Let us denote with ∂_i the neighbourhood of nodes adjacent to i in a given network, i.e. the set of nodes linked with i. The local clustering coefficient C_i is a property of node i counting the ratio of pairs in ∂_i that are connected with each other:

$$C_i = \frac{|\{(k,l) \in E \ for \ k,l \in \partial_i\}|}{\frac{|\partial_i|(|\partial_i| - 1)}{2}}.$$

In other words, the local clustering coefficient measures a tendency for a node's neighbours to be linked with each other. C_i ranges between 0 (no neighbours are linked) and 1 (all neighbours are linked) (Siew & Vitevitch 2019). Alternatively, C_i measures how much the neighbourhood of i resembles a complete network (where all nodes are adjacent to each other).

In Figure 4, the word "remark" has 3 neighbours that are all linked with each other, so $C_{remark} = 1$. Higher local clustering in phonological networks corresponded to degraded lexical identification, exemplified by the increased difficulty participants faced in accurately and swiftly recognizing and retrieving specific words during linguistic tasks (Chan & Vitevitch 2009). This was interpreted as activation accumulating along nodes in ∂_i and making it more difficult for i to stand out. For example in a phonological network with higher local clustering, participants were found to struggle to quickly distinguish between similar-sounding words due to the increased interference and competition among closely interconnected/clustered phonological representations (Castro et al. 2020).

In multiplex networks C_i cannot be generalised in a single way, depending on how links from different layers are counted (De Domenico 2022). Siew and Vitevitch (2019) aggregated orthographic and phonological similarities and found that a higher C_i corresponded to a facilitatory effect over spoken word recognition. In a multiplex network combining orthographic and phonological layers, a higher C_i value might indicate a more interconnected relationship between written and spoken representations of words, leading to enhanced facilitation in recognizing and processing those words in spoken form (Siew & Vitevitch 2019).

2.8. Connectedness, distance and closeness centrality in single-layer and multiplex networks

Network distance is the shortest number d of links connecting any two nodes (cf. Steyvers & Tenenbaum 2005). In Figure 4, "letter" and "phrase" are at distance $d_{letter,phrase} = 3$. If there is a path of any finite, non-zero, length between i and j then the latter are said to be connected. In simple networks, the largest set of connected nodes is also called the *largest connected component* (LCC). In directed networks, one has to consider directionality for navigating the network structure from one node to another. Two nodes might have a sequence of links between them (i.e. a path) but directionality might impede getting out from one node to access another along the path. If this happens, those nodes would be said to be weakly connected (Newman 2018). Otherwise, if directionality would still provide access from one node to another along any given path, those nodes would be said to be strongly connected (Newman 2018).

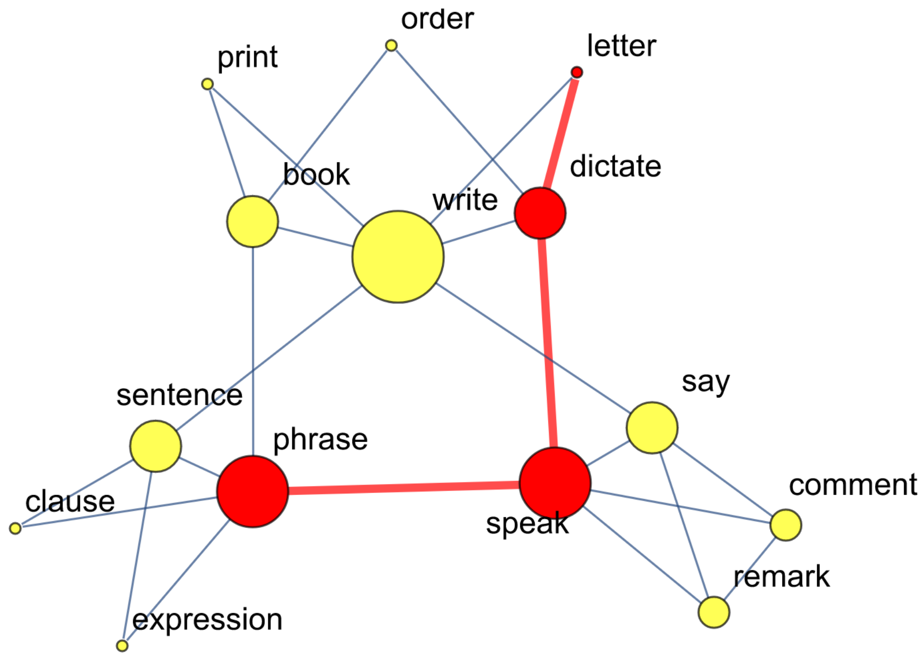

Figure 4: A cognitive network where nodes represent words and links indicate undirected free associations (i.e. two words were mentioned as cue and recall in a free association game). The shortest path between "letter" and "phrase" is highlighted in red. Node size is proportional to their closeness centrality.

In multiplex lexical networks, multiplex connectivity can exploit paths across all available layers such as connections between phonological and semantic layers within the same multiplex network (Stella et al. 2018). Consequently, multiplexity provides more shortcuts or opportunities for connecting any two nodes. Furthermore, replica nodes disconnected in one layer might be connected on another layer and thus, potentially, on the whole multiplex network too (see Fig. 2).

Cognitive network distance proves effective in capturing semantic relatedness and priming effects, as demonstrated by Kenett et al. (2017), Kumar et al. (2020), and Wulff et al. (2022). Kenett and colleagues (2017) found that in Hebrew, network distance in a free association network (where links indicate memory recall patterns) outperformed advanced models like latent semantic analysis when predicting data from a semantic relatedness judgement task. Kumar and colleagues (2020) replicated these findings in English and expanded the exploration of semantic distances to predict reaction times in priming tasks. They observed a linear increase in response latencies with network distance, revealing significant differences between prime-target cases $d_{pt} = 1$ and $d_{pt} = 4$. For reviews on this topic see (Kumar 2021; Wulff et al. 2022).

Notably, network distance has also illuminated variations in sequential lexical retrieval across different populations, including those with distinct traits such as task-quantified creativity levels

(Stella & Kenett 2019). Assessments like the verbal fluency task serve as valuable tools for exploring creativity and its relation to semantic networks (see also Section 3.2). In a verbal fluency task individuals are asked to generate words within specific constraints, e.g. naming as many animals as possible in two minutes (animal fluency task) (Goñi et al. 2011). Fluency tasks can be used in various forms, such as semantic fluency, where participants need to generate words belonging to a certain semantic category, or phonological fluency with the restriction that words need to begin with a particular letter. In semantic fluency tasks, individuals with higher creativity levels were found to produce recalls separated by longer network distances, on a multiplex network including almost 16k English words, compared to controls with lower creativity levels (Stella & Kenett 2019).

Whereas distance is a network property considering two nodes at a time, one could be interested also in assessing whether a single node is distant or close to all other nodes in a given connected component. In simple networks, this can be done by considering closeness centrality (Newman 2018). If i belongs to a connected component with N nodes, then the closeness of i is defined as:

$$c_i = \frac{1}{\langle d_{ij} \rangle} = \frac{N-1}{\sum_j d_{ij}},$$

where the normalisation $N - 1$ factor excludes contributions from self-loops with $d_{ii} = 0$.

Closeness centrality is a global network feature because it considers the whole layout of shortest paths between a node and its whole connected component. A shorter average network distance is supposed to be advantageous for mental navigation because it allows for quicker and more efficient access to information, facilitating faster information retrieval and processing (Hills et al. 2012; Dubossarsky et al. 2017; Stella et al. 2018; Levy et al. 2021). In an auditory lexical decision task, a higher closeness of words in phonological networks corresponded to shorter reaction times, highlighting a beneficial role of shorter network distance, thus higher closeness, over lexical decision (Goldstein & Vitevitch 2017).

In multiplex lexical networks, where there is no explicit cost for transitioning from one layer to another, network distance is defined as the shortest path length connecting any two nodes through links of any layer/colour (Stella et al. 2018). For instance, if “cat” and “crab” are connected by a direct link in the semantic layer (since both are animals) and another direct link in the phonological layer (due to their phonological similarity), the network distance between “cat” and “crab” is considered as 1, regardless of the layers involved in the connection. Castro and Stella (2019) showed how higher multiplex closeness centrality predicted better performance in picture naming by people with aphasia disorders. As already mentioned above, multiplexity can provide additional shortcuts, lowering down the distance between words (Quispe et al. 2021). Levy and colleagues showed that the links provided by a phonological network present substantially more shortcuts to free association networks than random expectation, thus, enhanced cognitive advantage over word processing from combining free associations and phonological similarities (Levy et al. 2021). Words with higher multiplex closeness centrality were found to be learned earlier by normative learning toddlers (Stella et al. 2017). This latter finding is in agreement with the preferential acquisition hypothesis in cognitive psychology (Hills et al. 2009; Sizemore et al. 2018; Borovsky et al. 2021; Cox & Haebig 2022): Children exhibit a tendency to preferentially acquire information that is more central or connected within the knowledge being transmitted in their environment (cf. Cox & Haebig 2022). Essentially, the preferential acquisition hypothesis suggests that the ease of access and interconnectedness of a concept within one's cognitive structure both positively influence the likelihood of that concept being learned earlier in the developmental process (Hills et al. 2009).

2.9. Beyond pairwise networks: Hypergraphs as cognitive models

Hypergraphs are generalisations of simple or multiplex networks where edges connect more than two nodes at the same time (Citraro et al. 2023; Citraro et al. 2023c). In k-regular hypergraphs, all hyperedges link exactly k different nodes. If k is not mentioned, then a hypergraph can feature hyperedges linking a variable amount of nodes.

Hypergraphs can be valuable models of the mental lexicon when dealing with subsequent recalls, like in continued free association games. In these tasks (cf. De Deyne et al. 2013), an individual reads a cue and then recall at most three responses related to it (e.g. reading “complex” might lead one to think about “system”, “emergent” and “more”). De Deyne and colleagues (2013) showed that building pairwise free association networks, i.e. featuring only links between the cue and its 3 responses, led to structural network measures that predicted a variety of lexical processing tasks better than alternative pairwise networks linking the cue and all responses with each other or considering only connection between the cue and the first response.

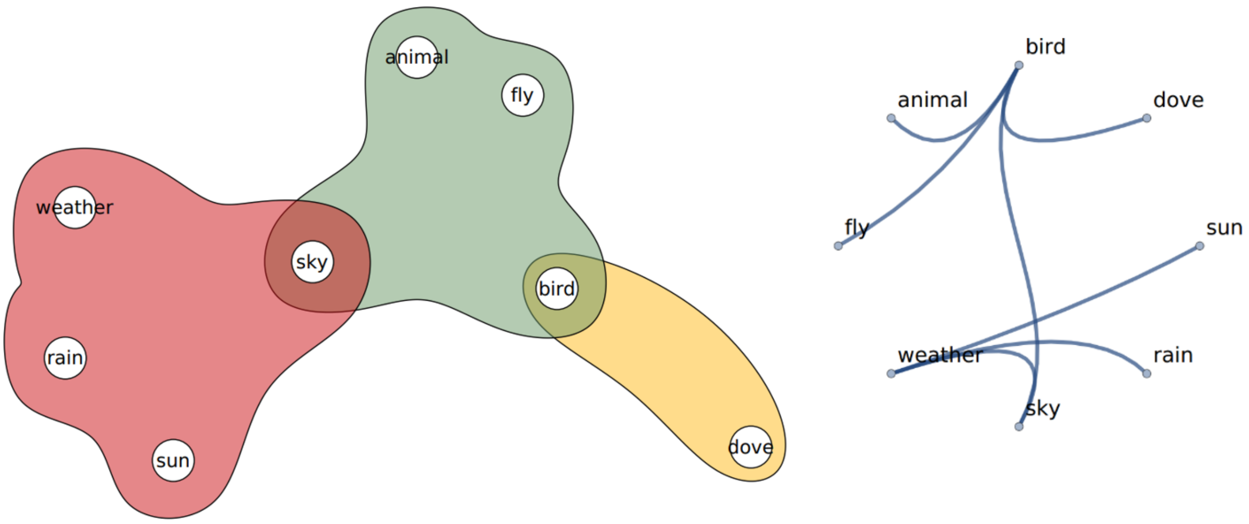

Figure 5: Cognitive hypergraph (left) and pairwise network (right) for the free association data relative to {{"dove", "bird"},{"bird","animal","fly","sky"},{"weather","sky","rain","sun"}}. Each set of responses is coded as a hyperlink but the first entry represents a cue, others are responses.

Revisiting free association data, Citraro and colleagues (2023; 2023b) introduced cognitive hypergraphs as a new idea at the interface of psychology and data science. Cognitive hypergraphs, where an hyperlink encapsulated both the cue and its responses, provided better models than pairwise networks in predicting psycholinguistic features like concreteness, a psycholinguistic norm indicating whether concepts are perceived as tangible or easy to picture (cf. Citraro et al. 2023) or predicting normative early word learning across self-reported norms and empirical norms in toddlers (cf. Citraro et al. 2023c). These findings underline that in some specific tasks relative to lexical acquisition and access, considering weak relationships between targets can be insightful. Hypergraphs provide more flexible tools for capturing conceptual relationships between more than two elements at a time and the above pioneering results indicate cognitive hypergraphs as a promising direction for future research in cognitive network science.

3. Datasets and software for building cognitive networks

Computational scientists interested in analysing cognitive networks should be aware of relevant datasets tailored to specific research interests. This section reviews both historical and more recent datasets, organised across different aspects of the mental lexicon. Python libraries like NetworkX (https://networkx.org/, Accessed: 12/12/23) and iGraph (https://igraph.org/, Accessed: 12/12/23) make it easy to import networks as edge lists, while providing a wide variety of pre-implemented network measures. Notice that iGraph works also in R and Mathematica. For R users, MuxViz is a powerful tool for multiplex network analysis and visualisation (De Domenico 2022). SemNA is another valuable R package that specialises in generating semantic networks out of verbal fluency data (Christensen & Kenett 2021).

3.1. Datasets and software for phonological networks

Phonological networks map sound similarities between words (Vitevitch 2008). These similarities can be identified through shared phonemes, rhyme patterns or similar phonetic features (Vitevitch & Mullin 2021). Building phonological networks thus requires phonological transcriptions. A widely used approach is to adopt the International Phonetic Alphabet (IPA) and consider IPA transcriptions as strings of characters with each character representing a phoneme (Turnbull 2021). This is advisable due to potential problems with orthographic writing, where the same sequence of letters may represent different sounds in different words for non-transparent languages (Aitchison 2012), e.g. the English word "meet" differing from its IPA transcription /miːt/. IPA also contains characters not representing phonemes directly but still encoding for the duration of other phonemes. Therefore, it is most common to use phonetic word transcriptions for generating phonological networks (Vitevitch 2008), including or excluding non-phoneme characters (Turnbull 2021). For "transparent languages", where there is a consistent mapping between phonemes and letters, phonological networks can be built out of orthographic word transcriptions (Martinčić-Ipšić et al. 2016). The "transparency" of these languages is constituted by these consistent rules for grapheme-phoneme mapping, e.g. learning that "ie" in German is always pronounced /iː/. For non-transparent languages, like English, ad-hoc datasets are required to build phonological networks. The command WordData[] in Mathematica contains phonological transcriptions for over 35,000 English words. A more extensive resource is the IPA-DICT project (https://github.com/open-dict-data/ipa-dict, Accessed: 03/04/22), which gathers IPA transcriptions for 23 languages in multiple formats (including tsv, JSON and csv). Data scientists might also consider procedural pipelines, which read an orthographic input and apply sequences of transformations to get phonological transcriptions. One of these pipelines is the Epitran project (Mortensen et al. 2018), which comes as an easy-to-install Python library and supports over 90 languages.

As mentioned above, the local topology of words in a phonological network can correspond to facilitative/inhibitive patterns in word retrieval, confusability or short-term retention (cf. Siew et al. 2019; Vitevitch & Mullin 2021). The R package *spreadR* simulates spreading activation over any single-layer network structure, including phonological ones, to capture patterns of lexical retrieval like false alarm rates and inhibited spoken word recognition (Siew 2019). The package can be loaded via CRAN and gets as an input a network topology, the starting node, a retention rate (how much activation should not spread across links at each timestep), a dampening factor, the initial activation level and how many steps should be executed. The package can output the whole spreading dynamics, i.e. sequences of activation levels across all timesteps for all nodes in the network. Extensively tested in phonological and semantic networks, *spreadR* represents a practical tool for harnessing how patterns of word recall or identification might emerge from the interplay of network structure and spreading dynamics (cf. Siew 2019).

3.2. Datasets and software for semantic networks

Semantic networks codify meaning similarities or memory-based relationships between words such as hierarchical relationships like “animal” and “mammal”, or associative links like “sun” and “day” (Kenett et al. 2018; Wulff et al. 2019; Valba et al. 2021). Dictionaries historically represented key data sources for building semantic networks (Miller 1995). However, Big Data and mega-studies are quickly enriching the landscape of semantic data by providing vast automatic datasets (Steyvers & Tenenbaum 2005; Kennington & Schlangen 2015; Kumar et al. 2020). In this tutorial paper, we will focus on: (i) dictionary-based data, and (ii) behavioural/psychological data.

For dictionary-based data, one of the largest corpora of semantic relationships is WordNet (Miller 1995), where words are organised according to sense and semantically related senses create *synsets*. For instance, “duck.n.01” refers to the animal with n.01 referring to a noun, whereas “duck.v.01” is a movement with v.01 referring to a verb. WordNet maps six semantic relationships:

1. **Synonyms**: Synonymous words that share an overlap in meaning in some contexts, e.g. “forest” and “woods”;
2. **Hypernyms**: Hypernyms represent superordinate generalisations, i.e. a concept representing a whole category of more specific concepts, like “bird” being a generalisation of “dove”;
3. **Hyponyms**: Hyponyms are specifications of broader categories, i.e. a concept being a specific instance of another concept representing a broader category, like “apple” being a specification of “fruit”.
4. **Antonyms**: Antonyms denote contraries, like “trust” and “distrust”.
5. **Meronyms**: Meronyms highlight a concept being a part of another, as seen in “wheel” being a part and, thus, a meronym of “car”.

6. **Holonyms**: Holonyms depict a concept including another, such as “cell phone” and “speaker”.

In Python, WordNet 3.0 is available via the *wordnet* package in NLTK (the Natural Language ToolKit, see https://www.nltk.org/howto/wordnet.html, Accessed: 12/12/23). WordNet is also available in R and, natively, in Mathematica through the WordData[] command. Through the packages, one can access synsets and use them as edge lists (Steyvers & Tenenbaum 2005).

Whereas dictionary-based data comes from lexical resources mostly curated by expert linguists, behavioural/psychological data include different types of semantic data coming from experimental setups.

Free association data allow us to examine how individuals recall information in relation to a specific task, with the resulting patterns of retrieval being coded as empirical data, i.e. requiring an experiment to be obtained rather than coming from dictionaries (De Deyne et al. 2013; Kenett et al. 2018; Wulff et al. 2022; Wulff et al. 2022b). As mentioned in Section 2.9, in the continued free association task a participant reads a cue (e.g. “pen”) and has to react with up to three responses that come to their mind as quickly as possible (e.g. “paper”, “writer”, “pencil”) (De Deyne et al. 2013; De Deyne et al. 2019). Free associations encode memory recalls, which are mostly semantic (Stella et al. 2017) but also multidimensional (De Deyne et al. 2019; Ciaglia et al. 2023): Free associations can come from phonological, orthographic, visual and other types of similarities between words. Currently, the Small World of Words (SWOW) project is the largest repository of free association data across 17 languages (https://smallworldofwords.org/en/project/home, Accessed: 12/12/23; cf. De Deyne et al. 2019). Unlike traditional dictionaries or thesauri, SWOW utilises word associations to uncover the meanings and centrality of words in the human mind. The SWOW project involves presenting participants with words and prompting them to provide spontaneous associations, i.e. to perform the continued free association task. SWOW represents a rich, large-scale dataset: Currently, SWOW counts over 700,000 links between 40,000 English words and it keeps growing (De Deyne et al. 2019). Free associations powered a wide variety of applications in the last few years, including the interplay between semantic memory and concepts related to beauty and wellbeing (Kenett et al. 2021), or suicide ideation (Teixeira et al. 2021), or personality traits (Kenett et al. 2018). Free association networks using SWOW enabled researchers in exploring also the intricacies of word relationships, including complex dimensions like individual differences (Wulff et al. 2022), creativity levels (Stella & Kenett, 2019), mental well-being (Fatima et al. 2021) and cognitive processes across ageing (Dubossarsky et al. 2017; Wulff et al. 2019; Wulff et al. 2022b). SWOW thus represents a promising dataset for next-generation models of cognition and human behaviour.

Other notable mentions of free association datasets include the University of South Florida Free Association norms - 70,000 links between 10,000 American English words/nodes (Nelson et al. 2004) - and the Edinburgh Associative Thesaurus - 325,000 links between 23,000 British English words/nodes (Wilson & Kiss, 1988).

Behavioural semantic data can come also as fluency lists (Rastelli et al. 2022; Fatima et al. 2021; Siew & Guru 2023), i.e. ordered sequences of words relative to a given category (as already mentioned in Section 2.8). For instance, in a category task, individuals produce fluency lists of as many words as possible from a given category within a time limit of 2 minutes (Goñi et al. 2011). The number of responses measures the so-called linguistic fluency and it has been used historically as a psychometric measurement of language skills (see also Zemla et al. 2020). However, fluency lists should be seen as the outcome of mental search and retrieval processes of a highly complex nature and potentially foraging/exploiting some network aspects of the mental lexicon (Hills & Kenett 2021). Cognitive networks built from fluency lists can go beyond simple fluency measures and unveil the structure among consecutive responses. How to assess the cognitive network structure embedded but not immediately apparent in fluency lists? There is no unique answer to this research question, however most studies on the topic tend to follow mostly two psychometric approaches: (i) consider fluency lists as ordered sequences of cognitive data where subsequent or co-occurring responses share stronger semantic similarities than concepts further apart in the sequence, or (ii) consider fluency lists as observed data coming from a latent cognitive network structure, which can be inferred via mathematical modelling.

Historically, the first approach was the first one to appear in the literature (cf. Goñi et al. 2011), inspired by word co-occurrences (Amancio 2015). Goñi and colleagues (2011) introduced a simple yet elegant co-occurrence Binomial filtering for tracing which concepts tended to co-occur more than randomly expected in sets of fluency lists (Goñi et al. 2011). Their approach was thus psychometric in that they considered a null model and filtered against spurious co-occurrences. When conducting the category fluency task on the topic “animals”, the resulting fluency networks were typically characterised such that animals sharing several semantic features (such as appearance or habitat) tended to be more tightly connected to each other than to other animals/nodes in the network.

Fluency data can also be considered as the outcome of random walks over latent cognitive structures. Hence, making semantic networks based on fluency data can be considered the outcome of an inference problem. This means that building such networks involves making mathematical inferences based on the fact that the observed fluency data is the outcome of a stochastic/random process, one observation among many possible ones. A tool that can be utilised for building fluency networks in this way is the package *SNAFU* in Python (https://github.com/AusterweilLab/snafu-py, Accessed: 04/12/23). *SNAFU* stands for *Semantic Network and Fluency Utility* and helps, among other functions, with counting clusters, perseverations and calculating word frequency (Zemla et al. 2020). This reconstruction approach created cognitive networks whose features automatically classified people affected by early Alzheimer's Disease (against healthy controls) with an accuracy of over 92% and an F1 score of 0.88 (Zemla & Austerweil 2019). These datasets and methodologies all open the way to exciting data-informed explorations of mental search processes.

An intermediate (set of) approach(es) between network inference and co-occurrence analysis is the recent *SemNA* package, which transforms fluency data in cognitive networks and provides accessible tools for network analysis with a comprehensive point-and-click interface (Christensen & Kenett 2021). This package incorporates several R modules, including SemNetDictionaries, SemNetCleaner, and SemNeT. These modules play integral roles in preprocessing, data cleaning, and semantic network generation. Notably, SemNA facilitates the generation of four distinct networks, each grounded in diverse network theories (Christensen & Kenett 2021). Figure 5 portrays a snippet of example code in R for using the SemNet package. In a first step, the verbal fluency data is prepared by using the *textcleaner* function for automatic spelling correction, identification and exclusion of inappropriate words. Next, group data is divided in order to generate semantic networks and calculate differences for each group, including statistical testing via t-tests and bootstrapping. Finally, the SemNetShiny() command opens an easily accessible point-and-click interface where statistical and graphical semantic network analyses are conducted.

```r

## Semantic Network Analysis with SemNet package (Christensen & Kenett, 2021)
#-----------------------------------------------------------------
#   Load packages
#-----------------------------------------------------------------
library(SemNetCleaner)
library(SemNeT)
library(shiny)

#-----------------------------------------------------------------
#   Preparing Verbal Fluency Data
#-----------------------------------------------------------------
vf.school.data <- read.data("VF_school.xlsx", header = TRUE)
vf.school.data.short <- vf.school.data[,-c(1:2, 4)]
vf.school_clean <- textcleaner(data = vf.school.data.short, miss = 99,
                               partBY = "row", dictionary = "Schule")
                               # Custom-made dictionaries can be used here.

#-----------------------------------------------------------------
#   Separating groups
#-----------------------------------------------------------------
group <- ifelse(vf.school.data$Group == "control", "control", "intervention")
vf.control <- vf.school_clean$responses$clean[which(group
                                                    == "control"),]
vf.intervention <- vf.school_clean$responses$clean[which(group
                                                         == "intervention"),]

write.csv(vf.control, file = "control.csv", row.names = TRUE)
write.csv(vf.intervention, file = "intervention.csv", row.names = TRUE)

#-----------------------------------------------------------------
#   Estimating and Analyzing Semantic Networks
#-----------------------------------------------------------------
SemNeTShiny()
# Launch SemNetShiny() to access the easy point-and-click interface.
```

Figure 5: Example code used within the SemNet package in R (Christensen & Kenett 2021). The package has an integrated module for preprocessing and cleaning raw fluency data. Dictionaries for British and American English are available and can be extended with custom-made dictionaries (see *dictionary = "NAME OF DICTIONARY"*). Lastly, by launching SemNetShiny(), one is forwarded to an easily accessible point-and-click interface where semantic network analyses are conducted.

Recent studies used fluency networks (built in either ways) that can unveil crucial differences in the way groups of individuals with different altered cognitive states (Rastelli et al. 2022), age (Cosgrove et al. 2021), education levels (Denervaud et al. 2021) or domain knowledge proficiency (Siew & Guru 2023) structured their semantic memories. These recent and intriguing results underline how giving structure to fluency data can unveil interesting cognitive and psychological patterns.

3.3. Datasets and software for syntactic networks

Syntactic networks explicate grammatical and meaning relationships between words in sentences (Ferrer i Cancho et al. 2004). For instance, "Love is weakness" syntactically specifies a property ("weakness") of a noun, "love", through a verb. This sentence might thus be represented as the undirected link (*love, weakness*). Extracting syntactic relationships is known as *parsing*. In the human mind, this complex task involves a variety of complex phenomena taking place in sequences that are partly unknown (cf. Doczi 2019) but include: breaking down sentences into their grammatical components, identifying the relationships between words and understanding the overall meaning conveyed by the arrangement of linguistic elements. In the context of syntactic analysis for research, the task of parsing has long been left to human coding (Carley 1993), which refers to the manual process of annotating and categorising linguistic elements and relationships between them within a given text. The time-consuming process of human coding has been revolutionised with the advent of soft computing models like the Stanford Universal Parser (Dozat et al. 2017) or spaCy's library (https://spacy.io/usage/linguistic-features#dependency-parse, Accessed: 12/12/23), which implement automatic syntactic parsing via Artificial Intelligence (AI).

Before automatic syntactic parsing, word co-occurrences have been a computationally advantageous proxy for extracting local syntactic relationships from text (Amancio 2015). Word-word co-occurrences identify which words in sentences tend to appear as adjacent or separated by l words. Co-occurrences can be discarded if below a threshold (Amancio 2015; Teixeira et al. 2021) or filtered against random expectation (Quispe et al. 2021). For a pedagogic tutorial about how to build word co-occurrence networks in Python, see the Semantic Brand Score Package (https://github.com/iandreafc/semanticbrandscore-demo, Accessed: 12/12/23), which was used to monitor brand perceptions online via co-occurrences in users' reviews (Colladon 2018). Word co-occurrences in adults' child-directed speech contributed to predicting early word learning (Stella et al. 2017) and differed in structure from adults' general speech (Cox & Haebig 2022).

While more computationally costly, automatic syntactic parsing captures non-local syntactic dependencies that would be lost with co-occurrences. Consider the sentence "Climate change is such a terrible, disastrous, problematic issue" in which "change" and "issue" are syntactically related but separated by three adjectives. Thus, co-occurrence networks with $l < 3$ would miss this syntactic relation due to the number of intervening specifications.

Capturing all syntactic relationships in a text means automating content mapping (Carley 1993) and operationalising the automatic constructions of sets of words all referring to a given target word x, i.e. the automatic construction of the semantic frame for x (Fillmore & Baker 2012). Both frame semantics and content mapping are frameworks capturing stances, frames, mindsets or perceptions, which in general terms all rely on specific ways of associating and combining ideas in communicative intentions (Stella 2022). According to the theory of frame semantics (Fillmore & Baker 2012), reconstructing the semantic frame of a word is enough to understand its meaning in a given text. Differently put, frame semantics indicates that the meaning of a word can be inferred from other words that "keep it company", i.e. syntactically related concepts.

Syntactic networks can, thus, reconstruct a set of stances, frames and even, to a limited extent, a fragment of the mindset of a text author, as shown with recent work about *textual forma mentis* networks (TFMNs, forma mentis meaning "mindset" in Latin) (Stella 2020). Differently from other NLP (Teixeira et al. 2021) or co-occurrence (Quispe et al. 2021) proxies to syntactic links, textual forma mentis networks enrich AI-detected syntactic links with cognitive/synonym relationships and affective/emotional norms to better represent the cognitive/affective content of texts. Importantly, TFMNs use an AI to perform syntactic parsing and then, on the resulting tree of syntactic dependencies, TFMNs link only pairs of words possessing meaning and at a distance lower than a threshold t on the syntactic tree, rather than in the original sentence. This difference with word co-occurrence networks crucially makes it possible for TFMNs to capture syntactic relationships between words not adjacent or close in a sentence and yet syntactically related.

Figure 6 reports a toy example of a TFMN: "failure" by itself is rated negatively by humans (cf. the Emotion Lexicon by Mohammad and Turney 2013). However, in the sentences "Failure and success are both sides of the same coin. Learn how to face failure." (see Fig. 6) it is syntactically linked with mostly positive concepts such as "learn" and "success". The TFMN approach thus unveils that "failure" in that sentence is framed along a positive connotation (i.e. overcoming failure). Textual forma mentis networks can be built through the EmoAtlaslibrary in Python (https://github.com/alfonsosemeraro/emoatlas, Accessed: 24/12/23) and they can be used to identify stances as well as emotional profiles of different texts. The package has to be downloaded from github and, once imported, enables the extraction of TFMN edge lists and the visualisation of specific semantic frames (like the one in Fig. 6, left) with the aid of colours and hierarchical edge bundling (i.e. a heuristic placing nodes on a circle and putting those sharing links closer). This granularity has been used in studies to highlight crucial differences in the ways COVID-19 vaccines were framed by mainstream media and alternative news outlets (Semeraro et al. 2022).

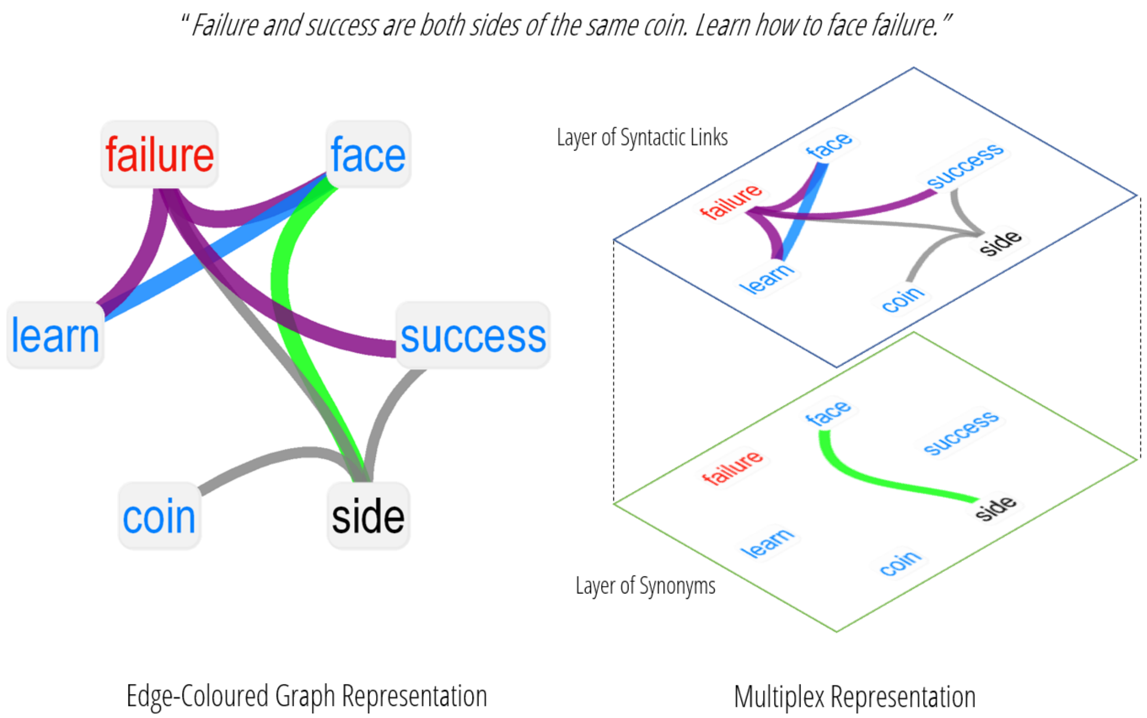

Figure 6: A textual forma mentis network (TFMN) for the text "Failure and success are both sides of the same coin. Learn how to face failure." Words are highlighted in red (black, cyan) if rated as negative (neutral, positive) in the Emotional Lexicon. Syntactic links are coloured as the words at their endpoints. Synonyms are in green and come from WordNet. TFMNs can be visualised as either edge-coloured graphs (left) or multilayer networks (right).

Another compelling option for generating text for syntactic analysis is the Woseco task, short for "word-sentence-construction". Originating as a teaching method developed by Haim and Aschauer (2022), the Woseco task aims to enhance the verbal creativity of pupils within the context of a specific school subject. In this task, an instructor provides a technical term from a broad category. Task participants have then to construct a factually correct sentence that not only includes the provided category but also incorporates a second technical term - of their choice - with which they will build another sentence, related to the category. The process is iterated several times (Haim & Aschauer 2022). Leveraging automated analysis tools like spaCy in Python (https://spacy.io/, Accessed: 24/11/23), it becomes possible to extract target words (typically nouns) from these sentences, building a network of syntactic dependencies that differs from TFMNs. This integration of Woseco into syntactic analysis not only provides a valuable resource for understanding language dynamics but can also enrich the exploration of verbal creativity in educational contexts.

Syntactic parsing can also be enhanced by entity recognition. Entity recognition involves identifying and classifying specific entities, such as names, locations, or other meaningful elements, within a given text or context. Enhancing syntactic parsing through entity recognition allows for a more detailed understanding of the relationships and roles of identified entities in the analysed text. The recent *netts* library (NETworks of Transcript Semantics, https://pypi.org/project/netts/, Accessed: 04/12/23) can disambiguate entities in texts, e.g. "Exeter – the University" vs "Exeter – the city", and draw syntactic relationships between clusters of them, such as "Exeter –> is a red brick –> University" (Nettekoven et al. 2022). Relying on CoreNLP and its recurrent neural network/matching models (https://stanfordnlp.github.io/CoreNLP/, Accessed: 04/12/23), *netts* is currently available in Python. When analysing transcripts produced by people affected by psychosis, *netts* unveiled key differences in network structure compared to healthy controls. This represents evidence that syntactic networks can enable a deeper understanding of ways of thinking as encoded also in texts written by clinical populations. The sentence structure and grammar of such clinical texts can provide insight into the thought processes of clinical populations which can be better understood through cognitive network science (Lydon-Staley et al. 2019; Parola et al. 2022).

4. Limitations and Future Work

Cognitive network science is deeply rooted in the mental lexicon metaphor and in associative knowledge modelling. Cognitive networks are powerful in giving structure to apparently unstructured data (e.g. texts, recalls, lexical tasks). They, thus, unveil patterns in the organisation, acquisition and framing of conceptual representations that cannot be obtained as easily with unstructured modelling frameworks, e.g. cognitive networks outperform latent semantic spaces in explaining similarity ratings (Kenett et al. 2017). However, there are also several limitations to the field that should be acknowledged, discussed and potentially relaxed via future research.

Firstly, whereas the mental lexicon is highly dynamic (Aitchison 2012; Zock 2020), most cognitive network models are currently static, i.e. the layout of conceptual associations does not change over time. However, pioneering studies with free associations (Dubossarsky et al. 2017; Wulff et al. 2019; Wulff et al. 2022b) and fluency data (Cosgrove et al. 2021) have shown that indeed semantic memory in the lexicon changes over time. Alterations in the mental lexicon can be due to various factors such as experience, learning and exposure to new information (Wulff et al. 2022). As individuals encounter new words and concepts, their mental lexicon dynamically adapts, incorporating new associations to reflect evolving semantic structures and relationships (Storkel 2002; Beckage et al. 2011). Next-generation cognitive network models should thus be able to account for time-evolving conceptual associations. Relevant applications could be the investigation of treatment effects (Veltri 2023) or other aspects of ageing (Wulff et al. 2022). A promising route would be the adoption of recent frameworks like stream graphs, where individual links can appear, persist and vanish over time, altering node centrality and network dynamics (cf. Citraro et al. 2021).

Secondly, the mental lexicon is multi-faceted (Zock 2020) and cognitive networks must encompass multiple aspects of knowledge at once. Multiplex lexical networks already account for multiple types of conceptual links (Stella et al. 2017; Stella et al. 2018). The combination of these various conceptual links highlights phenomena invisible to single-layer networks, e.g. distance-based mechanisms in predicting picture naming performance in people with aphasia (Castro & Stella 2019). However, these models are limited to mapping only the same set of nodes across layers, e.g. only words, whereas the mental lexicon might include elements like phonemes or sentences (Aitchison 2012), or even specific regions of the human brain activated by priming (Zaharchuk & Karuza 2021). Multilayer networks, i.e. generalisations of edge-coloured graphs where different sets of nodes can be linked across layers (De Domenico 2022), represent a valuable future direction for quantitative interpretations of psycholinguistic data. With some pioneering studies investigating the multilayer structure of language (Martinčić-Ipšić et al. 2016), future research should deploy these models for interpreting more psychological data.

Thirdly, whereas most studies work at group level (Castro & Siew 2020), e.g. comparing creativity between scientists and artists (Merseal et al. 2023) or across different age groups (Cosgrove et al. 2023), cognitive network models of individuals' mental lexica might provide a major modelling advancement. Cognitive networks generated on the level of individuals might be especially beneficial in digital ecosystems where user-level data is already available (Mokryn & Ben-Shoshan 2021), thus potentially mapping how individuals perceive, link and express ideas through online communicative intentions. The SWOW database (Small World of Words, https://smallworldofwords.org/en/project/home, Accessed: 24/11/23) recently collected free associations at individual level (Wulff et al. 2022; Wulff et al. 2022b). This enables a granularity that is crucial for pushing cognitive network science towards embracing individual assessment and variability, making cognitive networks interesting objects of investigations even for researchers already familiar with complex networks, e.g. from network psychometrics (Golino & Epskamp; Golino et al. 2020).

Fourthly, cognitive networks are not always the best model for explaining cognitive phenomena. Kumar and colleagues found that word embeddings obtained from the word2vec method provided results slightly better than network distances in explaining priming data (Kumar et al. 2020). However, the patterns reproduced by word2vec and cognitive networks differed, so that combining model results together could be more informative. When not directly superior, word embeddings might confirm results found through cognitive networks. Analysing narratives provided by clinical populations, the word embedding approaches of (Litovsky et al. 2022) and of (Parola et al. 2022) both identify abnormal patterns analogous to ones detected by syntactic networks in (Nettekoven et al. 2022) or described in (Lydon-Staley et al. 2019). Next-generation models should use vectorial and network aspects of the mental lexicon in synergy. A promising example is DASentimental (Fatima et al. 2021), which combines semantic network distance and word embeddings to predict levels of anxiety, depression and stress from combinations of emotional words. Using a cognitive network embedding technique, a multi-perceptron model was trained that demonstrated an ability to predict emotional distress levels in individuals with a performance comparable to that of human raters (cf. https://github.com/asrafaiz7/DASentimental, Accessed: 12/12/2023). Mixing cognitive networks with AI techniques might improve the explainability of the latter and power the creation of human-centred models where associative knowledge might improve not only AI performance but also foster novel methodological tools for computational social science and affine fields (Veltri 2023).

Last but not least, cognitive networks primarily examine the structure of conceptual connections, yet the mental lexicon includes a distributional dimension, representing words as numerical vectors (Kumar 2021; Litovsky et al. 2022). Addressing the challenge of reconciling these two aspects, the recent Feature Rich Multiplex Lexical Network (FERMULEX) framework integrates both the vector and network elements of associative knowledge. This approach reveals patterns in early sentence production by toddlers that remain unnoticed by models solely focused on either network or vector representations (Citraro et al. 2022). While FERMULEX considers only language-based features (e.g. frequency, length, polysemy), future models should account also for word-level task-based features such as concreteness (Citraro et al. 2023), specificity or how context-specific concepts are perceived (Bolognesi & Caselli 2023) or even latency - how easily concepts can be recalled or identified (Kumar et al. 2020; Litovsky et al. 2022). Extending network measures to feature-rich networks might also provide additional insights over the dichotomous network/vector nature of the mental lexicon. As an idealised system, the mental lexicon influences knowledge processing in ways mediated by vector similarities in some phenomena and by network structure in others (cf. Aitchison 2012; Zock 2020; Kovacs et al. 2021; Hills & Kenett 2021; Citraro et al. 2022). This underlines the need for computational investigations of such complex cognitive systems through a *multiverse* approach (Parola et al. 2022), where multiple models are compared, combined and reconciled.

5. Conclusions

This tutorial paper has explored recent advancements in cognitive networks, emphasising their role as models for understanding human cognition and behaviour. Their applications range from deciphering cognitive processes across visual, auditory, and semantic tasks in diverse populations to predicting cognitive development, decline, and performance in both clinical and healthy contexts. Additionally, cognitive networks contribute to the reconstruction of semantic framing within texts and media. As a useful tool for modelling human behaviour, the field of cognitive networks stands poised for growth, guided by meticulous statistical modelling, collaborative synergy with other interpretable frameworks, and the availability of rich datasets and powerful software packages spanning various tasks and contexts.

References

1. Abrams, L., & Davis, D. K. (2016). The tip-of-the-tongue phenomenon. *Cognition, language and aging*, 13-54.
2. Aitchison, J. (2012). Words in the mind: An introduction to the mental lexicon. *John Wiley & Sons.*
3. Alstott, J., Bullmore, E., & Plenz, D. (2014). powerlaw: a Python package for analysis of heavy-tailed distributions. *PloS one*, *9*(1), e85777.
4. Amancio, D. R. (2015). A complex network approach to stylometry. *PloS one, 10(8), e0136076.*
5. Aschauer, W., Haim, K., & Weber, C. (2022). A contribution to scientific creativity: A validation study measuring divergent problem solving ability. *Creativity Research Journal*, *34*(2), 195-212.
6. Battiston, F., Nicosia, V., & Latora, V. (2017). The new challenges of multiplex networks: Measures and models. *The European Physical Journal Special Topics*, *226*, 401-416.
7. Baker, O., Montefinese, M., Castro, N., & Stella, M. (2023). Multiplex lexical networks and artificial intelligence unravel cognitive patterns of picture naming in people with anomic aphasia. *Cognitive Systems Research*, *79*, 43-54.
8. Beaty, R. E., Kenett, Y. N., Christensen, A. P., Rosenberg, M. D., Benedek, M., Chen, Q., ... & Silvia, P. J. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences*, *115*(5), 1087-1092.
9. Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PloS one, 6(5), e19348.*
10. Bolognesi, M. M., & Caselli, T. (2023). Specificity ratings for Italian data. *Behavior Research Methods*, *55*(7), 3531-3548.
11. Borovsky, A., Thal, D., & Leonard, L. B. (2021). Moving towards accurate and early prediction of language delay with network science and machine learning approaches. *Scientific Reports, 11*(1), 1-12.
12. Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, *39*, 510-526.
13. Carley, K. (1993). Coding choices for textual analysis: A comparison of content analysis and map analysis. *Sociological methodology*, 75-126.

14. Castro, N., & Siew, C. S. (2020). Contributions of modern network science to the cognitive sciences: Revisiting research spirals of representation and process. *Proceedings of the Royal Society A, 476*(2238), 20190825.
15. Castro, N., & Stella, M. (2019). The multiplex structure of the mental lexicon influences picture naming in people with aphasia. *Journal of Complex Networks, 7*(6), 913-931.
16. Castro, N., Stella, M., & Siew, C. S. (2020). Quantifying the interplay of semantics and phonology during failures of word retrieval by people with aphasia using a multiplex lexical network. *Cognitive Science*, *44*(9), e12881.
17. Champagnat, N., Deaconu, M., Lejay, A., Navet, N., & Boukherouaa, S. (2013). An empirical analysis of heavy-tails behavior of financial data: The case for power laws.hal-00851429
18. Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1934.
19. Christensen, A. P., & Kenett, Y. N. (2021). Semantic network analysis (SemNA): A tutorial on preprocessing, estimating, and analyzing semantic networks. *Psychological Methods, 28(4), 860–879.*
20. Ciaglia, F., Stella, M., & Kennington, C. (2023). Investigating preferential acquisition and attachment in early word learning through cognitive, visual and latent multiplex lexical networks. *Physica A: Statistical Mechanics and its Applications*, *612*, 128468.
21. Citraro, S., De Deyne, S., Stella, M., & Rossetti, G. (2023). Towards hypergraph cognitive networks as feature-rich models of knowledge. *EPJ Data Sci.* 12, 31.
22. Citraro, S., Milli, L., Cazabet, R., & Rossetti, G. (2022). Δ-Conformity: multi-scale node assortativity in feature-rich stream graphs. *International Journal of Data Science and Analytics*, 1-12.
23. Citraro, S., Vitevitch, M. S., Stella, M., & Rossetti, G. (2023b). Feature-rich multiplex lexical networks reveal mental strategies of early language learning. *Scientific Reports*, *13*(1), 1474.
24. Citraro, S., Warner-Willich, J., Battiston, F., Siew, C. S., Rossetti, G., & Stella, M. (2023c). Hypergraph models of the mental lexicon capture greater information than pairwise networks for predicting language learning. *New Ideas in Psychology*, *71*, 101034.
25. Colladon, A. F. (2018). The semantic brand score. Journal of Business Research, 88, 150-160.
26. Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review, 82(6), 407*.
27. Cosgrove, A. L., Kenett, Y. N., Beaty, R. E., & Diaz, M. T. (2021). Quantifying flexibility in thought: The resiliency of semantic networks differs across the lifespan. *Cognition, 211*, 104631.
28. Cosgrove, A. L., Beaty, R. E., Diaz, M. T., & Kenett, Y. N. (2023). Age differences in semantic network structure: Acquiring knowledge shapes semantic memory. *Psychology and aging, 38(2),* 87–102.
29. Cox, C. R., & Haebig, E. (2022). Child-oriented word associations improve models of early word learning. *Behavior Research Methods,* 1-22.
30. De Arruda, H. F., Silva, F. N., Costa, L. D. F., & Amancio, D. R. (2017). Knowledge acquisition: A Complex networks approach. *Information Sciences*, *421*, 154-166.
31. De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods, 45*(2), 480-498.
32. De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods, 51(3)*, 987-1006.

33. De Domenico, M. (2022). Multilayer Networks: Analysis and Visualization, Springer Cham, Geneva, Switzerland.
34. Denervaud, S., Christensen, A. P., Kenett, Y., & Beaty, R. E. (2021). Education shapes the structure of semantic memory and impacts creative thinking. *npj Science of Learning, 6*(1), 1-7.
35. Dóczi, B. (2019). An overview of conceptual models and theories of lexical representation in the mental lexicon. *The Routledge handbook of vocabulary studies*, 46-65.
36. Dóczi, B., & Kormos, J. (2015). Longitudinal developments in vocabulary knowledge and lexical organization. Oxford University Press.
37. Dozat, T., Qi, P., & Manning, C. D. (2017, August). Stanford's graph-based neural dependency parser at the conll 2017 shared task. In Proceedings of the CoNLL 2017 shared task: Multilingual parsing from raw text to universal dependencies (pp. 20-30).
38. Dubossarsky, H., De Deyne, S., & Hills, T. T. (2017). Quantifying the structure of free association networks across the life span. *Developmental psychology, 53*(8), 1560.
39. Fatima, A., Li, Y., Hills, T. T., & Stella, M. (2021). DASentimental: Detecting Depression, Anxiety, and Stress in Texts via Emotional Recall, Cognitive Networks, and Machine Learning. *Big Data and Cognitive Computing, 5*(4), 77.
40. Fillmore, C. J., & Baker, C. F. (2012). A frames approach to semantic analysis. In The Oxford handbook of linguistic analysis (pp. 313-339). Oxford University Press.
41. Goldstein, R., & Vitevitch, M. S. (2017). The influence of closeness centrality on lexical processing. *Frontiers in psychology, 8*, 1683.
42. Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS one, 12*(6), e0174035.
43. Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., ... & Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*(3), 292.
44. Goñi, J., Arrondo, G., Sepulcre, J., Martincorena, I., Vélez de Mendizábal, N., Corominas-Murtra, B., ... & Villoslada, P. (2011). The semantic organization of the animal category: evidence from semantic verbal fluency and network theory. *Cognitive processing, 12*(2), 183-196.
45. Haim, K., & Aschauer, W. (2022). Fostering Scientific Creativity in the Classroom: The Concept of Flex-Based Learning. *International Journal of Learning, Teaching and Educational Research*, *21*(3), 196-230.
46. Hills, T. T., & Kenett, Y. N. (2021). Networks of the Mind: How Can Network Science Elucidate Our Understanding of Cognition?. *Topics in Cognitive Science, 14*(1), 189-208.
47. Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review, 119*(2), 431.
48. Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition?. *Psychological Science, 20*(6), 729-739.
49. i Cancho, R. F., Solé, R. V., & Köhler, R. (2004). Patterns in syntactic dependency networks. *Physical Review E, 69*(5), 051915.
50. Kenett, Y. N., Levi, E., Anaki, D., & Faust, M. (2017). The semantic distance task: Quantifying semantic distance with semantic network path length. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(9), 1470.
51. Kenett, Y. N., Levy, O., Kenett, D. Y., Stanley, H. E., Faust, M., & Havlin, S. (2018). Flexibility of thought in high creative individuals represented by percolation analysis. *Proceedings of the National Academy of Sciences, 115*(5), 867-872.

52. Kenett, Y. N., Ungar, L., & Chatterjee, A. (2021). Beauty and Wellness in the Semantic Memory of the Beholder. *Frontiers in Psychology, 12*, 696507.
53. Kennington, C., & Schlangen, D. (2015, July). Simple learning and compositional application of perceptually grounded word meanings for incremental reference resolution. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (pp. 292-301).
54. Koponen, I. T. (2021). Systemic states of spreading activation in describing associative knowledge networks II: generalisations with fractional graph Laplacians and q-adjacency kernels. *Systems, 9*(2), 22.
55. Kovács, L., Bóta, A., Hajdu, L., & Krész, M. (2021). Networks in the mind–what communities reveal about the structure of the lexicon. *Open Linguistics, 7*(1), 181-199.
56. Kumar, A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletin & Review, 28*(1), 40-80.
57. Kumar, A. A., Balota, D. A., & Steyvers, M. (2020). Distant connectivity and multiple-step priming in large-scale semantic networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(12), 2261.
58. Levy, O., Kenett, Y. N., Oxenberg, O., Castro, N., De Deyne, S., Vitevitch, M. S., & Havlin, S. (2021). Unveiling the nature of interaction between semantics and phonology in lexical access based on multilayer networks. *Scientific Reports, 11*(1), 1-14.
59. Litovsky, C. P., Finley, A. M., Zuckerman, B., Sayers, M., Schoenhard, J. A., Kenett, Y. N., & Reilly, J. (2022). Semantic flow and its relation to controlled semantic retrieval deficits in the narrative production of people with aphasia. *Neuropsychologia, 170*, 108235.
60. Lydon-Staley, D. M., Barnett, I., Satterthwaite, T. D., & Bassett, D. S. (2019). Digital phenotyping for psychiatry: Accommodating data and theory with network science methodologies. *Current opinion in biomedical engineering, 9*, 8-13.
61. Martinčić-Ipšić, S., Margan, D., & Meštrović, A. (2016). Multilayer network of language: A unified framework for structural analysis of linguistic subsystems. *Physica A: Statistical Mechanics and its Applications, 457*, 117-128.
62. Merseal, H. M., Luchini, S., Kenett, Y. N., Knudsen, K., Bilder, R. M., & Beaty, R. E. (2023). Free association ability distinguishes highly creative artists from scientists: Findings from the Big-C Project. *Psychology of Aesthetics, Creativity, and the Arts*.
63. Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM, 38*(11), 39-41.
64. Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational intelligence, 29*(3), 436-465.
65. Mokryn, O., & Ben-Shoshan, H. (2021). Domain-based Latent Personal Analysis and its use for impersonation detection in social media. *User Modeling and User-Adapted Interaction, 31*(4), 785-828.
66. Mortensen, D. R., Dalmia, S., & Littell, P. (2018, May). Epitran: Precision G2P for many languages. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).
67. Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402-407.
68. Newman, M. (2018). *Networks: An Introduction*. Oxford university press, London (UK).
69. Nettekoven, C. R., Diederen, K., Giles, O., Duncan, H., Stenson, I., Olah, J., ... & McGuire, P. (2023). Semantic speech networks linked to formal thought disorder in early psychosis. *Schizophrenia Bulletin*, *49*(Supplement_2), S142-S152.

70. Parola, A., Lin, J. M., Simonsen, A., Bliksted, V., Zhou, Y., Wang, H., ... & Fusaroli, R. (2023). Speech disturbances in schizophrenia: Assessing cross-linguistic generalizability of NLP automated measures of coherence. *Schizophrenia Research*, *259*, 59-70.
71. Quispe, L. V., Tohalino, J. A., & Amancio, D. R. (2021). Using virtual edges to improve the discriminability of co-occurrence text networks. *Physica A: Statistical Mechanics and its Applications, 562*, 125344.
72. Rastelli, C., Greco, A., Kenett, Y. N., Finocchiaro, C., & De Pisapia, N. (2022). Simulated visual hallucinations in virtual reality enhance cognitive flexibility. *Scientific Reports, 12*(1), 1-14.
73. Semeraro, A., Vilella, S., Ruffo, G., & Stella, M. (2022). Emotional profiling and cognitive networks unravel how mainstream and alternative press framed AstraZeneca, Pfizer and COVID-19 vaccination campaigns. *Scientific Reports 12*, 14445.
74. Siew, C. S. (2019). spreadr: An R package to simulate spreading activation in a network. *Behavior Research Methods, 51*(2), 910-929.
75. Siew, C. S., & Guru, A. (2023). Investigating the network structure of domain-specific knowledge using the semantic fluency task. *Memory & Cognition*, *51*(3), 623-646.
76. Siew, C. S., & Vitevitch, M. S. (2019). The phonographic language network: Using network science to investigate the phonological and orthographic similarity structure of language. *Journal of Experimental Psychology: General, 148*(3), 475.
77. Siew, C. S., Wulff, D. U., Beckage, N. M., & Kenett, Y. N. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity, 2019*, 2108423.
78. Sizemore, A. E., Karuza, E. A., Giusti, C., & Bassett, D. S. (2018). Knowledge gaps in the early growth of semantic feature networks. *Nature Human Behaviour, 2*(9), 682-692.
79. Stanghellini, F., Perinelli, E., Lombardi, L., & Stella, M. (2023). Introducing semantic loadings: Bridging cognitive network science with network psychometrics for understanding depression, anxiety and stress. *PsyArXiv, rdv58.*
80. Stella, M. (2020). Text-mining forma mentis networks reconstruct public perception of the STEM gender gap in social media. *PeerJ Computer Science, 6*, e295.
81. Stella, M. (2020b). Multiplex networks quantify robustness of the mental lexicon to catastrophic concept failures, aphasic degradation and ageing. *Physica A: Statistical Mechanics and its Applications, 554,* 124382.
82. Stella, M. (2022). Cognitive network science for understanding online social cognitions: A brief review. *Topics in Cognitive Science, 14*(1), 143-162.
83. Stella, M., Citraro, S., Rossetti, G., Marinazzo, D., Kenett, Y. N., & Vitevitch, M. S. (2022). Cognitive modelling with multilayer networks: Insights, advancements and future challenges. *arXiv preprint arXiv:2210.00500*.
84. Stella, M., & Kenett, Y. N. (2019). Viability in multiplex lexical networks and machine learning characterizes human creativity. *Big Data and Cognitive Computing, 3*(3), 45.
85. Stella, M., Beckage, N. M., & Brede, M. (2017). Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific Reports, 7*(1), 1-10.
86. Stella, M., Beckage, N. M., Brede, M., & De Domenico, M. (2018). Multiplex model of mental lexicon reveals explosive learning in humans. *Scientific Reports, 8*(1), 1-11.
87. Stella, M., Vitevitch, M. S., & Botta, F. (2022). Cognitive Networks Extract Insights on COVID-19 Vaccines from English and Italian Popular Tweets: Anticipation, Logistics, Conspiracy and Loss of Trust. *Big Data and Cognitive Computing*, *6*(2), 52.
88. Steyvers, M., & Tenenbaum, J. B. (2005). The large‐scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41-78.

89. Storkel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of child Language*, *29*(2), 251-274.
90. Teixeira, A. S., Talaga, S., Swanson, T. J., & Stella, M. (2021). Revealing semantic and emotional structure of suicide notes with cognitive network science. *Scientific Reports, 11*(1), 1-15.
91. Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic inquiry*, *8*(3), 505-520.
92. Tulving, E. (1972). Episodic and semantic memory. *Organization of memory*, *1*(381-403), 1.
93. Turnbull, R. (2021). Graph-theoretic Properties of the Class of Phonological Neighbourhood Networks. In Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics (pp. 233-240).
94. Valba, O., Gorsky, A., Nechaev, S., & Tamm, M. (2021). Analysis of English free association network reveals mechanisms of efficient solution of Remote Association Tests. *PloS one, 16*(4), e0248986.
95. Veltri, G. A. (2023). Harnessing heterogeneity in behavioural research using computational social science. *Behavioural Public Policy*, 1-18.
96. Vigliocco, G., Ponari, M., & Norbury, C. (2018). Learning and processing abstract words and concepts: Insights from typical and atypical development. *Topics in cognitive science*, *10*(3), 533-549.
97. Vitevitch, M. S. (2008). What Can Graph Theory Tell Us About Word Learning and Lexical Retrieval?. *Journal of speech, language, and hearing research, 51*(2), 408.
98. Vitevitch, M. S., & Mullin, G. J. (2021). What Do Cognitive Networks Do? Simulations of Spoken Word Recognition Using the Cognitive Network Science Approach. *Brain Sciences, 11*(12), 1628.
99. Wilson, M., & Kiss, G. (1988). EAT: the Edinburgh associative corpus. Oxford Text Archive Core Collection.
100. Wulff, D. U., Aeschbach, S., De Deyne, S., & Mata, R. (2022). Data from the myswow proof-of-concept study: Linking individual semantic networks and cognitive performance. *Journal of Open Psychology Data*, *10*(1).
101. Wulff, D. U., De Deyne, S., Aeschbach, S., & Mata, R. (2022b). Using Network Science to Understand the Aging Lexicon: Linking Individuals' Experience, Semantic Networks, and Cognitive Performance. *Topics in Cognitive Science*, *14*(1), 93-110.
102. Wulff, D. U., De Deyne, S., Jones, M. N., Mata, R., & Aging Lexicon Consortium. (2019). New perspectives on the aging lexicon. *Trends in Cognitive Sciences, 23*(8), 686-698.
103. Zaharchuk, H. A., & Karuza, E. A. (2021). Multilayer networks: An untapped tool for understanding bilingual neurocognition. *Brain and Language, 220*, 104977.
104. Zemla, J. C., & Austerweil, J. L. (2019). Analyzing knowledge retrieval impairments associated with Alzheimer's disease using network analyses. *Complexity, 2019*.
105. Zemla, J. C., Cao, K., Mueller, K. D., & Austerweil, J. L. (2020). SNAFU: The semantic network and fluency utility. *Behavior Research Methods, 52*(4), 1681-1699.
106. Zock, M. (2020). AI at the Crossroads of NLP and Neurosciences. *Journal of Cognitive Science, 21*(1), 1-14.