INVESTIGATIONES LINGUISTICAE VOL. XLV, 2021
©INSTITUTE OF ETHNOINGUISTICS ADAM MICKIEWICZ UNIVERSITY
AL. NIEPODLEGŁOŚCI 4, 60-874, POZNAŃ – POLAND

Growing Networks – Modelling the Growth of Word Association Networks for Hungarian and English

László Kovács
EÖTVÖS LORÁND UNIVERSITY
BUDAPEST, HUNGARY
kovacs.laszlo@sek.elte.hu

Katalin Orosz
EÖTVÖS LORÁND UNIVERSITY
BUDAPEST, HUNGARY
oroszk@hal.elte.hu

Peter Pollner
EÖTVÖS LORÁND UNIVERSITY
BUDAPEST, HUNGARY
pollner@angel.elte.hu

Abstract

In the new era of information and communication technology, the representation of information is of increasing importance. Knowing how words are connected to each other in the mind and what processes facilitate the creation of connections could result in better optimized applications, e.g. in computer aided education or in search engines.

This paper models the growth process of a word association database with an algorithm. We present the network structure of word associations for an agglutinative language and compare it with the network of English word associations. Using the real-world data so obtained, we create a model that reproduces the main features of the observed growth process and show the evolution of the network. The model describes the growth of the word association data as a mixture of a topic based process and a random process.

The model makes it possible to gain insight into the overall processes which are responsible for creating an interconnected mental lexicon.

Keywords: complex networks, semantic networks, word association, growing network model

1. Introduction

Networks and networked structures are extremely widespread. They are found and analyzed both in and across disciplines (Barabási 2016; Barthelémy 2011; Easley and Kleinberg 2010; Menczer et al. 2020). Although network metaphors in linguistics are not new, cf. the networks of Quillian (1968), the spreading activation by Collins and Loftus (1975) the semantic networks of Figueroa et al. (1976), the cognitive representation model of Bybee (1995), or the PDP models of Rogers and McClelland (2004), networked structures have become a focus of research again just recently, as networks and characteristics of networks are described with methods routing in physics (cf. Mehler et al., 2016; Siew et al. 2019; Siew and Vitevitch 2019; Vitevitch 2020).

Although networks are not new to linguistics, the notion "network" was used in linguistics previously rather as a metaphor to describe certain characteristics, such as the mental lexicon, while today they are real constructs of linguistic data which can be researched extensively by methods of network science (Vitevitch 2020). In networks, nodes (e.g. words) and links (e.g. associative connections) exist. The number, strength and distribution of the links reveals the structure of networks (cf. Barabási 2016).

Semantic networks are extensively analyzed with these new methods. There are several ways of constructing a semantic network. One widely used method is connecting the words in the same sentence up to a certain distance with a text corpus (Ferrer i Cancho and Solé 2001; Li et al. 2012). Another method is presented by Motter et al. (2002), where the connection between two words can be established if they express similar concepts. Networks can also be constructed from thesauri, where words are connected to the semantic categories in which they fall into (Steyvers and Tenenbaum 2010). A possibility for creating a semantic network is also provided by WordNet, where word forms and word meanings are linked according to different types of relations (Fellbaum 1998; Miller 1995) or the FrameNet initated by Charles Fillmore, where more than 10.000 lexical units are collected based on their semantic valences in frames (Ruppehofer et al. 2016).

Another possibility is to use word association data for creating a semantic network. In word association experiments a stimulus word is presented to participants, who have to name that which is primed by the stimulus. Several experiment designs exist and the found associative structures were analyzed in many contexts in the 20th century (Cramer 1968; Kent and Rosanoff 1910; Kiss et al. 1973; Postman and Keppel 1970).

When a network is created from word association data, the nodes are words, while the links between the words are presented when a word primed another word. The networks created can be directed (stimulus → primed word) and weighted (how many participants named the primed word to the stimulus). Networks created from word association data can contain not only semantic, but also e.g. phonetic connections.

The analysis of the mental lexicon is experiencing a new renaissance due to methodological progress in recent years. Today, on one hand it is easier to collect a large amount of data for word associations with web-based applications (see e.g. Gravino et al 2012). On the other hand, new methods are available to analyze the mental lexicon as network characteristics found in other natural or artificial networks can be applied to investigate the mental lexicon.

The investigation of network structures of the lexicon can reveal hierarchical structures or identify more densely connected regions and contribute to a better understanding of the mental lexicon by researching syntactic dependency (De Deyne et al. 2016), seeing multiplex connections in the lexicon (Stella 2019; Stella et al. 2017) or helping to understand the overall structure of the lexicon (Vitevitch et al., 2014). Despite the differences in network constructions, word association databases show similarities in many aspects such as small-world properties and scale-free structures (cf. De Deyne and Storms 2008a, 2008b; Ferrer i Cancho and Solé 2001; Gravino et al. 2012; Jung et al. 2010; Li et al. 2012; Motter et al. 2012; Steyvers and Tenenbaum 2005).

In this paper, we present semantic networks created from free association experiments for English and for Hungarian. We interpret the word association process in terms of the theory of Quillian (1968). That is, the semantic memory consists of entries and attribute values can be assigned to each entry that defines the entry–value relationships. Hence each attribute value is a word as well and values can be considered as an entry as well. Accordingly, a word association network may be conceived of as a special type of tagged network where words are objects and tags at the same time (Cattuto et al. 2007). Therefore, a word-word relationship can also be considered as a word-tag relationship. Tag statistics have been analysed from a viewpoint of taxonomy in Tibély et al. (2012). The distribution of tags is greatly influenced by the hierarchical relationships among the tags.

We show that a special type of tagged networks, namely word association networks, are built up from a mixture of subprocesses, which connect words to each other either randomly or by semantic similarity. We compare the resulting structures for English and Hungarian.

2. Word association experiments

We present results from a word association experiment not only for English, but also for an agglutinative language, Hungarian. Agglutinative languages have some special rules and a much higher variation in word forms than do non-agglutinative languages. In agglutinative languages morphosyntactical elements adjoin the word stem (for a general overview cf. Rounds 2001, for syntactical implications cf. É. Kiss 2002).

The Hungarian Word Associations (HWA) data collection began in 2008. A custom-designed web page (ConnectYourMind, in Hungarian Agykapocs) has been launched to collect the associations of volunteers who have registered at the page (Kovács 2013, Kovács – Orosz – Pollner 2021).

The association data were collected in two ways. First, all participants were presented with the same set of cue words – 134 cue words – in the same order. After the first 134 words the software choose cue words randomly from the database, from the earlier obtained responses (for details see Kovács 2013). Participants had to type the first response word that came to their mind for the cue word. The time and the date of the response along with the registration ID of the respondent are stored in a database. For the responses no lemmatization was performed. The raw response data have been carefully cleaned of misspelled words. The resulting dataset in the analyzed form contains about 16,000 words and 43,000 different associations collected from 700 participants.

Having compared our dataset with data from an English word association experiment, we chose the publicly available South Florida Free Word Association dataset (Nelson et al. 1998). Collection of the Florida word association data began in the 1970s. More than 72,000 different associations were collected as responses for about 5000 cue words from more than 6000 participants during the long-term and tedious data collection procedure. The data collection process on paper and pencil lasted several years. For the research we used a subset of the South Florida Database (Palla et al. 2005).

3. Model

The emergence of a semantic network can be modelled by a growing process. In online word association experiments, different cue words are coming up sequentially for the participants, hence, more and more words and connections are formed in the network. The process is similar to tagging, where objects may get different labels or tags, like photo-tags in Flickr or webpage-tags on the site del.icio.us. However, in tagging experiments the resulting network is always bipartite (links point from the object to its labels), in word association the network consists of unique types of nodes (links point from words to other words. It should be noted that although we use the term 'word' for the response, depending on the experiment the response might be an expression or a sequence of words as well).

Free word association experiments build up the network through the first neighbours of the already existing network nodes: the aggregation of edges starts from some initial source nodes and constantly broadens the list of possible source nodes with the associated words obtained. In the process, activation takes place, by which the cue word triggers a search mechanism in the respondents mind. We assume that for such searching processes three principle important strategies can be found: broadcast search, random walk and degree-biased strategy (Barrat et al. 2008). Broadcast search proceeds through the neighbours of each intermediate step node until the desired piece of information is obtained at the target node. If a random walk is performed, the walker steps from the current node to a randomly chosen neighbour. In the case of the degree-biased strategy the search is carried out by stepping towards the neighbours by a large degree. Degree-biased search is also known as the preferential attachment rule in general network growth models (Barabási and Albert 1999).

For semantic networks an early preferential attachment based network growth model was assumed by Dorogovtsev and Mendes (2001). Quayle et al. (2006) identify and combine two types of preferential rules, one for node degree and one for node similarity. Masucci and Rodgers (2006) also use two different preferential attachments. They modified the Dorogovtsev-Mendes model by distinguishing between local and global preferential attachment. Another approach to preferential attachment is defined by Hébert-Dufresne et al. (2011). They emphasize that the associations between words are due to semantic similarities, since each word is a member of semantic categories and the attachment of new nodes to the network is based on the preferential attachment by the number of nodes in that category.

Here we work with a model depicted in Figure 1 which distinguishes between random associations and associations due to semantic similarity. Random associations arise through two random processes (Fig 1a, Choice between two possible process). One process is based on a general random preferential rule, where words that are associated to many different cue words are more likely to be chosen as a target than words targeted by fewer cue words (Fig 1b, Random target choice, Degree preferential). The other process uses an unbiased random rule generating links between nodes with different semantic categories (Fig 1b, Random target choice, Unbiased). For semantic similarity associations we grouped the words into separate categories, where the members of a category were considered as semantically similar words. When a new word (association) is semantically similar to the cue, then the link between them connects words of the same category (Fig 1b, Similarity target choice). We assumed that the structural preferential rule of Hébert-

Dufresne et al. (2011) is valid, where large categories (categories with many words) tend to acquire new words more often than small categories. However, in our model each node gets exactly one category, overlap between categories is not allowed (see Discussion). The schematic overview of the network growth strategy is illustrated in Figure 1.

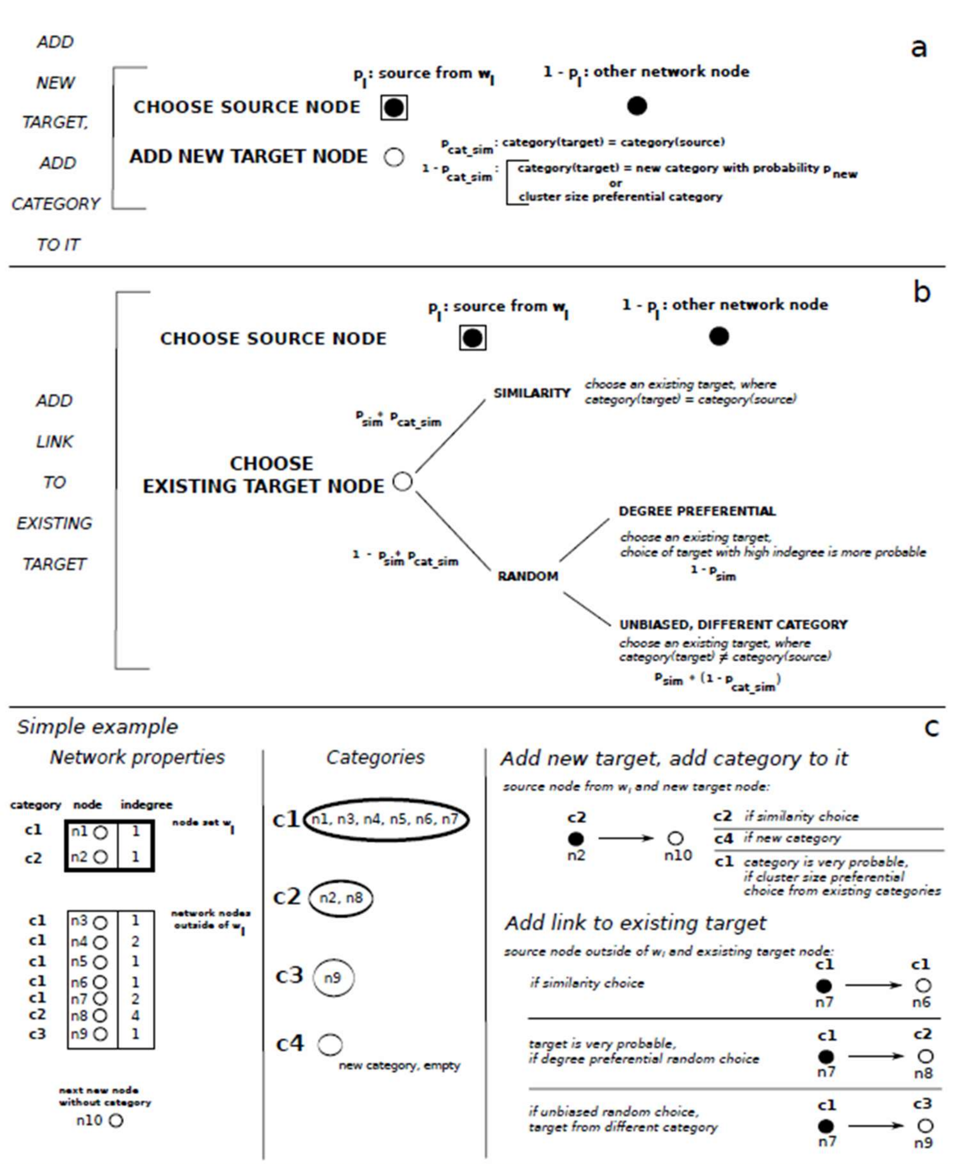

Figure 1: Schematic illustration of the model algorithm. a., adding a new node with a new edge, and adding the new node to a category b., adding a new edge between existing nodes c., simple example demonstrating the agorithm. w_I indicates the set of fixed cue words.

The resulting network of the model consists of unique type of nodes (words) connected by directed links which point from the cue word (source) to the response word (target). The network is weighted and the weight corresponds to the number of occurrences of the source → target pair.

In the model the network grows as follows: during the growth process edges and nodes are created in the network and each of the nodes is classified into a category. The categories are subsequently associated to the nodes as labels and they may be considered as a topic (a semantic field) the given word belongs to. As mentioned above, we have defined the categories as separated groups, therefore the categories form a non-overlapping partition of the network.

Initially a certain number of network nodes are selected as fixed cue words: nodes number 1,2,…,I. This set of words w_I will be handled specially in the process (Fig 1c, left most column).

In each step, after adding a new node with a single directed edge (the target of the edge is the new node), m edges are generated among the existing network nodes (Fig 1c, right most column). If a source → target pair already exists, then the weight of the directed edge will be increased. Self-loops are always excluded. New nodes are connected either to the special set w_I or to any other nodes. The special set is chosen for connection at a p_I rate (Fig 1a). For the Hungarian network this rate is changed during the growth process. At the start the p_I is initiated to 0.75 and later it decreases as the size of the network grows. In the large network size limit the p_I reaches an asymptotic value of 0.5. For the Florida network p_I is set to 1.

If a new node is attached to the network, it gets a category ID either by similarity or by random preferential choice. The category ID is attached by similarity at a constant rate $p_{cat\ sim}$ (Fig 1a). In this case the new node is classified into the category of its neighbour. In the other case when the category ID is attached by the preferential rule, the label of the new node is chosen randomly from a pool (Fig 1c right most column). The choice among the labels from the pool is random, and those labels are preferred that have been attached already to a larger number of nodes. The pool contains always a new category ID also, which is not attached to any node (Fig 1c middle column). This new category ID is selected with a constant probability, p_{new}.

New edges among the existing nodes are placed according to analogous rules as nodes are labelled into categories: two nodes are connected with a new link either by category similarity or by random choice. The probability for being connected by similarity is set to $p_{sim} * p_{cat\ sim}$. Similarity here refers to semantic similarity. In our model we do not, however, define semantic similarity – we were interested on the growth process itself.

If the new edge is placed by the random mechanism, the choice of the target node will be either biased by the indegree or unbiased. A biased indegree means that higher indegree nodes are more likely chosen as the target of the edge (words already connected to many words are likely to be connected to new words; as when somebody has a lot of friends, they are likely to make new friends easily). If it is unbiased, the target is selected from a pool of nodes containing nodes with category other than the category of the source node; that is each other node has the same probability to be connected to the source node.

The network is grown until the number of nodes reaches the desired value.

4. Results

In our analysis we will look at different network properties. Networks can be directed or undirected. In undirected networks, the existence of a connection between the nodes is important while in the case of a directed network the direction of the connection is also relevant. For example, in the case of a road network, roads connect settlements. The direction is not important, it is important that the road exists. In a citation network, direction is important as a scientific paper can only cite papers which already exist and cannot cite a paper which will be written ten years later. In this case, the network is a directed network. In a network, each node has a degree, which describes how many connections a given node has. In the case of directed networks, nodes have in- and out-degrees: indegree is the number of incoming, outdegree the number of outgoing connections. Scale-free networks are networks, in which the distribution of connections is described by the power law, which means that there are some highly connected nodes in the networks, there are less nodes with fewer connections and that most nodes have just a very few connections. Cluster size refers to the number of nodes that belong to the same cluster, where a cluster is a highly connected subpart of the network. Shortest paths describe the size of the networks: they show how each node can be reached from each other node by navigating through the existing connections and finding the shortest possible path.

We compared the network structure of the Hungarian word associations with the English dataset. The large scale structures of the word association networks in English and in Hungarian are similar. The indegree distributions of both the English and the Hungarian networks have scale-free behaviour and the outdegree distributions do not follow power-law decay. Cluster size distributions qualitatively fit to a power-law. Measured and model curves are highly similar for the indegree distributions at both networks. In the word list with the highest indegree nodes, we can find many equivalent concepts for the English and the Hungarian networks (Table 1). On the other hand, the maximum indegree is higher and the maximum outdegree is lower for the Florida network than for the Hungarian network. There is a second peak in the outdegree distribution of the Hungarian network. The largest Florida cluster (141 nodes) is approximately eight times smaller than the largest HWA cluster (1113 nodes). The shortest path length tends to be higher in the Hungarian association network. The maximum shortest path length of HWA and Florida networks are 21 and 11, respectively. We show some examples from the Hungarian and the English networks in Figure 2. The upper two subfigures demonstrate one subgraph from the Hungarian and one subgraph from the English network (Figure 2a and 2b). One of the longest paths (diameters) is shown in Figure 2c for the Hungarian and for the English network. We have highlighted the words that occur in both networks indicating their universal importance and illustrate similar universality with Table 1, where the most connected words are listed. Again, we see words that occur both in the Hungarian and in the English network in similar central position.

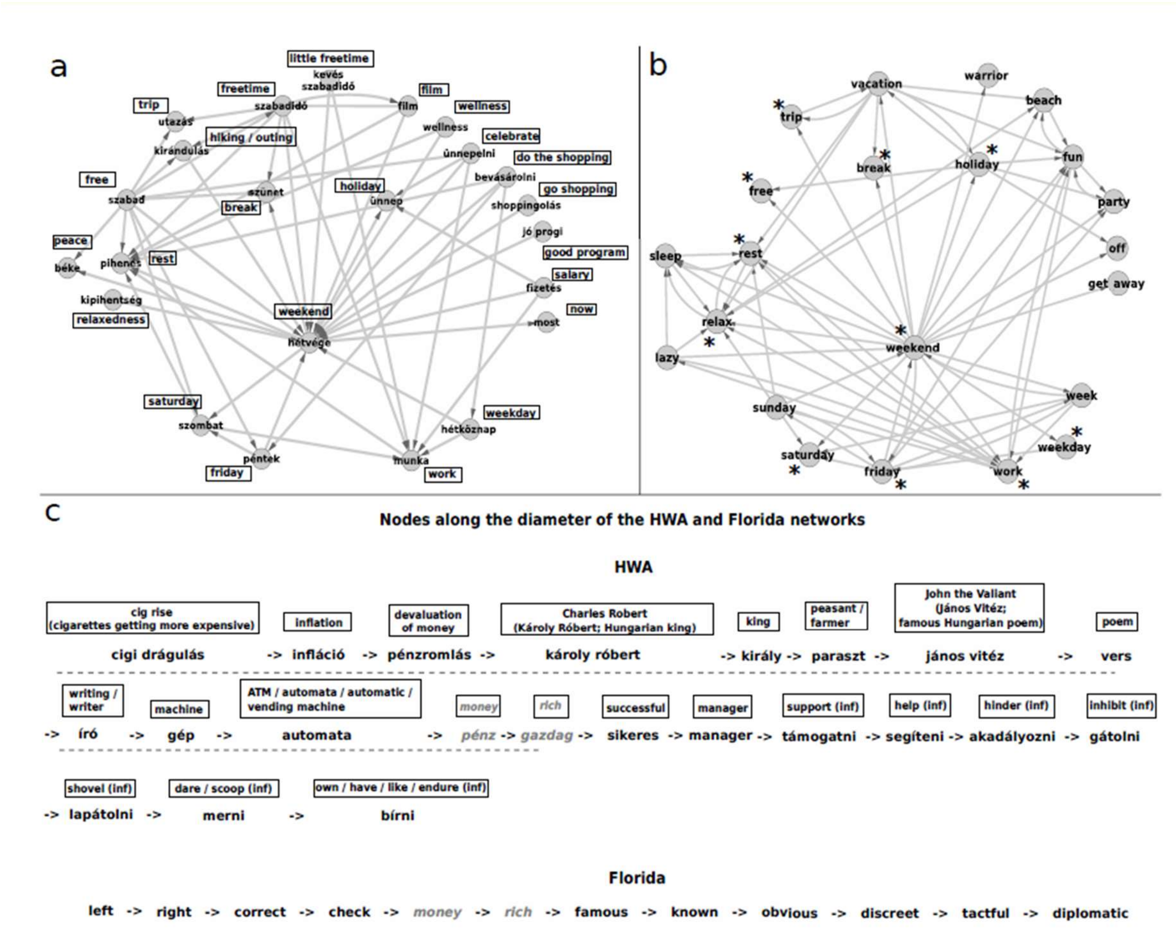

Figure 2: Examples from the HWA and Florida networks.
Translations of the Hungarian words are given in the boxes. a., a small section of the HWA network, b., a small section of the Florida network. Nodes marked by asterisk have a counterpart with a similar meaning in the HWA network. c., nodes along the diameter of the HWA and Florida networks. Similar concepts are indicated with a grey italic font.

	HWA	Florida
*	bad (rossz)	bad
*	car (autó)	car
		food
*	good (jó)	good
		house
	man (ember)	
	many/much (sok)	
*	money (pénz)	money
*	water (víz)	water
*	work (munka)	work

Table 1: Words with the highest indegrees in the HWA and Florida networks in alphabetical order. Words with equivalent meanings are marked by asterisk.

Besides comparing the Hungarian and English word association networks we used our model to measure the predictability of the association process. The model distinguished two basic processes: the network grows either by similarities or by random choices. Among the several parameters of the network we have defined the p_{sim} and $p_{cat\ sim}$ probabilities. They indicate how deterministic the growth process is. If these probabilities are close to 1, the network is very predictable, since the connections and category memberships are based on topic similarities. If these probabilities are small, the network is totally random.

We have fitted the model parameters to the network data by minimizing an error function. The error function was defined by the squared sum of differences in the distribution functions of the most important network parameters: the indegree, outdegree, category size and the shortest path length distributions. For the Hungarian network we have the possibility to compare the evolution of the real network and the evolution of the model as the network is growing step by step.

The similarity, when two nodes belong to the same category (two words are in the same topic), is a key concept in our model. Therefore we have compared the statistical properties of categories in the model with the categories in the real data. While in the model, similarity is defined through the generated labels, it was extracted from the real network data with the clustering of the network nodes. The categories of the English and Hungarian word association networks were determined by network community finding algorithms and words within a community were assigned to a category. We used the Cfinder (Palla et al. 2005) algorithm for finding the core categories. Since the networks were not dense enough and several nodes were outside of any communities, we have extended the communities by the Chinese Whispers clustering (CW) algorithm (Biemann 2006). Here, one node may belong to only one cluster and all of the network nodes are put into one of the clusters. CW clustering gave us clusters of words, where words in the same cluster tend to be related to the same topic. Our model requires the categories of the initial node set as input data which were extracted by running CW clustering on a subnetwork constructed from links amongst the initial cue words of the measured network. The number of within-cluster and inter-cluster edges are very close in the model and in the real networks, that validates our parameter fitting (Table 2).

	HWA	Model	Florida	Model
number of nodes	16562	16637 ± 0	10617	10520 ± 0
number of links	43702	45100 ± 86	72172	71697 ± 92
number of within-clust. links	21827	20038 ± 107	17292	17419 ± 103
number of inter-clust. links	21875	25062 ± 101	54880	54278 ± 121

Table 2: Network properties of the measured and the corresponding model networks. Values coming from the model are averaged for 20 model networks.

Figure 3 and Figure 4 illustrate our results for the HWA and Figure 5 for the Florida network. Figure 3 and Figure 5 a, b, c show the degree and the cluster size distributions, Figure 3 d and Figure 5 d are the shortest path length distributions for the HWA and Florida measured and the corresponding model networks. Each measured network is compared to 20 independent model calculations.
The model parameters for HWA and Florida model networks are shown in Table 3.

	HWA model	Florida model
I	134	5017
m	3	16
pcat sim	0.8	0.6
psim	0.15	0.35
pnew	0.2	0.2

Table 3: Model parameters for HWA and Florida model networks.

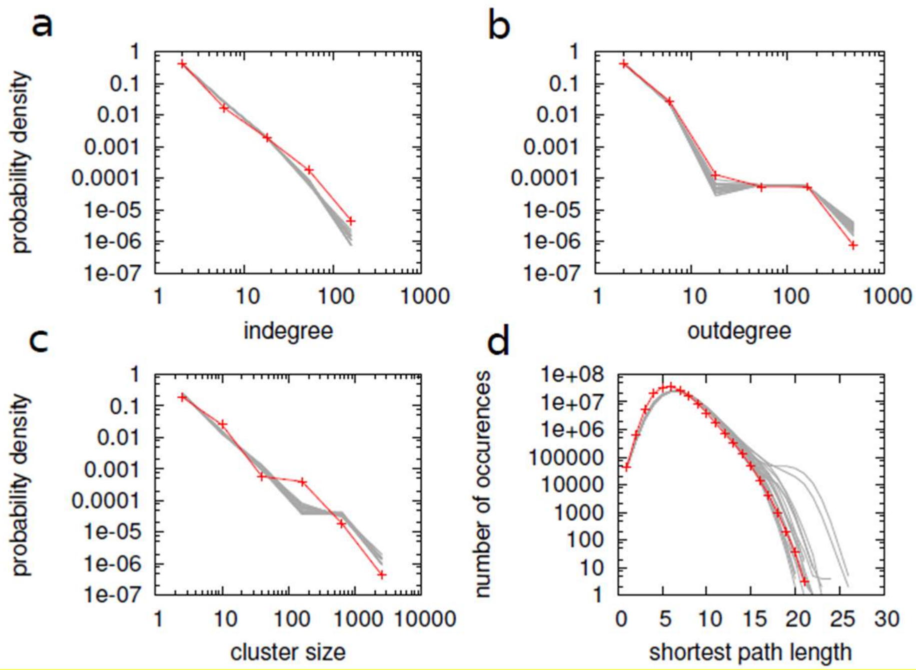

Figure 3: Model results compared to the HWA network: a., indegree distribution b., outdegree distribution c., cluster size distribution d., shortest path length distribution. The line with plus symbols belongs to the measurement and the other lines show the model calculations. Probability density is the normalized frequency of the data points. For better visibility data points are aggregated into bins.

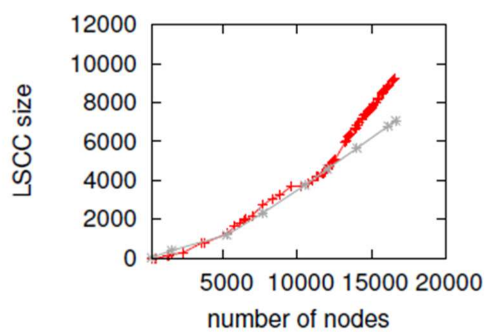

Figure 4: Model results compared to the HWA network: LSCC size evolution. The line with plus symbols belongs to the measurement and the other, with asterisk symbols, shows the average of the model calculations.

The HWA growth was analysed through the size of the largest strongly connected component (LSCC) in Figure 4. Our model results follow the growth obtained from the experiment up to 12000 nodes. Above this, the model growth is a bit slower than the measured one. In the case of the Florida network, we are able to calculate the LSCC size only for the final state of the growing network. Florida LSCC consists of 4845 nodes. This value is 4961 on average from the model calculations.

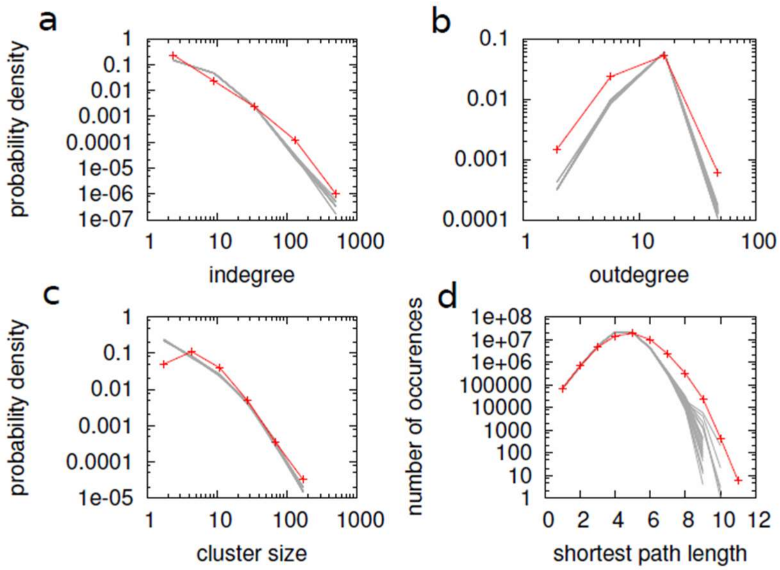

Figure 5: Model results compared to the Florida network: a., indegree distribution b., outdegree distribution c., cluster size distribution d., shortest path length distribution. The line with plus symbols belongs to the measurement and the other lines show the model calculations. Probability density is the normalized frequency of the data points. For better visibility data points are aggregated into bins.

The best fit of the model yield an estimate for the p_{sim} and $p_{cat-sim}$ probabilities. In both networks studied, we found, that the growth process involves a non-negligible random part. For the Hungarian network the random process takes part about 70% in the network evolution, and for the English network we found 77% randomness.

5. Discussion

Although the global network properties are similar for the two languages, we see several differences in the details that could be explained by structural differences between the two languages. Some differences in network characteristics could be explained by the structure of the Hungarian database (and of Hungarian as an agglutinative language): in the Hungarian database respondents were allowed to put words in every possible form into the database (e.g. also inflected, conjugated forms). Inflected forms can have the results, that the given form triggers one given association. Just one example: the Hungarian word házba means 'into the house'. When the inflected form házba is presented, a likely association will be bemenni or menni 'go into' or 'go'; while the form ház (house) triggers associations like kert ('garden') or lakás ('flat'). There is a difference then between associative structures of inflected and uninflected forms. In the English database however the form 'house' is used, and since it was not possible, as far as we know, to have multiple words as answers, the unit 'into the house' will not be part of the database; therefore the database contains the associative structure around 'house', but not around 'into the house'; while the Hungarian database may have both structures. This structural difference – coming from Hungarian as an agglutinative language depicted by the actual association network by allowing inflected forms to enter the database – could indicate a locally smaller density, be responsible for longer paths and for larger connected components.

The differences between the network structures may, however, also be explained by other factors such as the time scale in which the two databases were formed, by the method of data collection (paper vs. online), by the initial word lists or by cultural differences. Since however no other comparable Hungarian word association database exists, the assumptions cannot be confirmed or verified on other Hungarian data.

We must also note, that in the model each node was part of exactly one category: which puts a limitation on the study. We choose this model to keep the model as simple as possible: by adding possibilities for nodes to belong to more than one category (e.g. allowing polysemy), the model would be much more complex. Allowing polysemy would result in a more connected network, where polysemous words would function as bridges between different parts of the network decreasing the length of shortest paths. This is a research direction however, which must be considered when creating new growth models for word association databases.

6. Conclusion

We have presented a word association network for an agglutinative language. We have compared the main network properties of the Hungarian and English association networks and found similarities in the global structures: both languages have an association network with a small-world property, scale-free indegree and cluster size distribution.

We have presented a network growth model that follows similarity and random rules. The model parameters were fitted by minimizing the error in the distribution of the main network properties as degree, category size and shortest path length distributions. We also compared the number of within-cluster and inter-cluster links as well and found that the best fitted models must have a dominant random component.

Word associations can be interpreted also as a tagging process, during which words are tagged with other words. Therefore, our result hints for general tagging processes, that tags are attached to nodes not only by similarity processes, but there is an important random process as well.

Acknowledgements

We would like to thank Tamás Vicsek and Gergely Palla for their useful advice. This research was partially supported by the Hungarian National Research, Development and Innovation Office (grant no. K 128780). The founders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Barabási, A.-L. 2016: *Network Science*. Cambridge: Cambridge University Press.

Barabási, A.-L. and Albert, R. 1999: Emergence of scaling in random networks. *Science* 286(5439): 509–512.

Barrat, A., M. Barthlemy and Vespignani, A. 2008: *Dynamical Processes on Complex Networks*. Cambridge: Cambridge University Press.

Barthélemy, M. 2011: Spatial networks. *Physics Reports* 499(1): 1–101.

Biemann, C. 2006: Chinese whispers: an efficient graph clustering algorithm and its application to natural language processing problems. In: *Proceedings of the First Workshop on Graph Based Methods for Natural Language Processing*. New York: Association for Computational Linguistics. 73–80.

Bybee, J. 1995: Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5): 425–455.

Cattuto, C., Loreto, V. and Pietronero, L. 2007: Semiotic dynamics and collaborative tagging. *Proceedings of the National Academy of Sciences* 104(5): 1461–1464.

Collins, A. M. and Loftus, E. F. 1975: A Spreading-Activation Theory of Semantic Processing". *Psychological Review* 82(6): 407–428.

Cramer, P. 1968: *Word Association*. New York: Academic Press.

De Deyne, S. and Storms, G. 2008a: Word Associations: Norms for 1,424 Dutch words in a continous task. *Behavior Research Methods* 40(1): 198–205.

De Deyne, S. and Storms, G. 2008b: Word associations: Network and semantic properties. *Behavior Research Methods* 40(1): 213–231.

De Deyne, S., Verheyen, S. and Storms, G. 2016: Structure and Organization of the Mental Lexicon: A Network Approach Derived from Syntactic Dependency Relations and Word Associations. In: Mehler, A. et al. (eds.) *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*. Berlin–Heidelberg: Springer. 47–79.

Dorogovtsev, S. N., and Mendes, J. F. F. 2001: Language as an evolving word web. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268(1485): 2603–2606.

Easley, D. and Kleinberg, J. 2010: *Networks, Crowds and Markets. Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press.

É. Kiss, K. 2002: *The Syntax of Hungarian*. Cambridge: Cambridge University Press.

Fellbaum, C. (ed.) 1998: *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press.

Ferrer i Cancho, R. and Solé, R. V. 2001: The small world of human language, Proceedings of the Royal Society of London. *Series B: Biological Sciences* 268(1482): 2261–2265.

Figueroa, G. J., González, E. G. and Solís, V. M. 1976: An approach to the problem of meaning: Semantic networks. *Journal of Psycholinguistic Research* 5(2): 107–115.

Gravino, P. et al. 2012: Complex structures and semantics in free word association. *Advances in Complex Systems* 15(3-4): 1250054.

Hébert-Dufresne, L. et al. 2011: Structural preferential attachment: Network organization beyond the link. *Physical Review Letters* 107(15): 158702.

Jung, J., Na, L. and Akama, H. 2010: Network Analysis of Korean Word Associations. In: *Proceedings of the NAACL HLT 2010 First Workshop on Computational Neurolinguistics*. Los Angeles: Association for Computational Linguistics. 27–35.

Kent, G. H., and Rosanoff, A. J. 1910. A study of association in insanity. *American Journal of Insanity* 67(1-2): 37–96; 317–390.
Kiss, G. R. et al. 1973: An associative thesaurus of English and its computer analysis. In: Aitken, A. J., Bailey R. W. and Hamilton-Smith, N. (eds.) *The Computer and Literary Studies*. Edinburgh: Edinburgh University Press. 153–165.
Kovács, L. 2013: *Fogalmi rendszerek és lexikai hálózatok a mentális lexikonban*. Budapest. Tinta.
Kovács L., Orosz K. and Pollner P. 2021: Networks in the Mental Lexicon – Contributions from Hungarian. *Glottotheory* 12(2) (to be appear)
Li, J. et al. 2012: Chinese lexical networks: The structure, function and formation. *Physica A: Statistical Mechanics and its Applications* 391(21): 5254–5263.
Masucci, A. P. and Rodgers, G. J. 2006: Network properties of written human language. *Physical Review E* 74(2): 026102.
Mehler, A. et al. (eds.) 2016: *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*. Berlin–Heidelberg: Springer.
Menczer, F., Fortunato, S. and Davis, C. A. 2020: *A First Course in Network Science*. Cambridge: Cambridge University Press.
Miller, G. A. 1995: WordNet: a lexical database for English. *Communications of the ACM 38* (11): 39–41.
Motter, A. E. et al. 2002: Topology of the conceptual network of language. *Physical Review E* 65(6): 065102.
Nelson, D. L., McEvoy, C. L. and Schreiber, T. A. 1998: *The University of South Florida word association, rhyme, and word fragment norms*. Retrieved from http://w3.usf.edu/FreeAssociation/Intro.html
Palla, G. et al. 2005: Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435: 814–818.
Postman, L. and Keppel, G. (eds.) 1970: *Norms of Word Association*. New York: Academic Press.
Quayle, A. P., Siddiqui, A. S. and Jones, S. J. M. 2006: Modeling network growth with assortative mixing. *The European Physical Journal B-Condensed Matter and Complex Systems* 50(4): 617–630.
Quillian, M. R. 1968: Semantic memory. In: Minsky, M. (ed.) *Semantic Information Processing*. Cambridge: MIT Press. 227–270.
Rounds, C. H. 2001: *Hungarian: An Essential Grammar*. London: Routledge.
Rogers, T. T.–McClelland, J. L. 2004: *Semantic Cognition*. Cambridge: MIT Press.
Ruppenhofer, J. et al. 2016: *FrameNet II: Extended Theory and Practice*. Available online at https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf [Accessed 31. January 2022.]
Siew, C. S. Q. et al. 2019: Cognitive Network Science: A Review of Research on Cognition through the Lens of Network Representations, Processes, and Dynamics. *Complexity* 2019(5915): 1–24.
Siew, C. S. Q. and 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): 475–500.
Stella, M. 2019: Modelling Early Word Acquisition through Multiplex Lexical Networks and Machine Learning. *Big Data and Cognitive Computing* 3(1): 10.

Stella, M., Beckage, N. M. and Brede, M. 2017: Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific Reports* 7: 46730.
Steyvers, M., and Tenenbaum J. B. 2010: The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive science* 29(1): 41–78.
Tibély, G. et al. 2012: Ontologies and tag-statistics. *New Journal of Physics* 14(5): 053009.
Vitevitch, M. S. et al. 2014: Using complex networks to understand the mental lexicon. *Yearbook of the Poznan Linguistic Meeting* 1(1): 119–138.
Vitevitch, M. S. 2020: Introduction. In: Vitevitch, M. S. (ed.) *Network Science in Cognitive Psychology*. New York – London: Routledge. 1–9.