Knowledge transfer process by language networks

Fóris Ágota[*†]- Koltai Dénes[‡]- Kozma László[§]

HU ISSN 1418-7108: HEJ Manuscript no.: LIN-081029-B

Abstract

The aim of the paper is to examine the role of networks in optimizing the storage, transfer and use of information today, and to determine the role of language networks in the process of knowledge transfer. It is a widely accepted fact that network relationships are present in nature and society, and within the latter, in production, services, education, and in the operation of every other field. Knowledge of the working of networks makes it possible to design better systems and to establish the optimal conditions of operation. This implies the possibility that knowledge transfer process could be established on networks studies.

Az előadás célja annak vizsgálata, hogy hogyan jelenik meg napjainkban az információ tárolásának, átadásának és felhasználásának optimalizálásában a hálózat szerepe, és a tudásátadás folyamatában milyen szerepet játszanak a nyelvi hálózatok. Elfogadott tény, hogy a természet és a társadalom, ezen belül a termelés, igazgatás, szolgáltatás, oktatás és minden más terület működésében jelen vannak a hálózati kapcsolatok. A hálózatok működését leíró ismeretek lehetővé teszik bármely rendszer jobb tervezését, a működés optimális feltételeinek a kialakítását. Ennek a lehetőségnek a megvalósítására kapott eredmények felvetik annak lehetőségét, hogy a tudásátadási folyamat kialakítása a hálózatokra vonatkozó ismeretek alapján történjen meg.

Keywords: information, knowledge transfer process, language network, scale-free network

1 Introduction

It was recognised only a few years ago that almost every process in the world takes place in a network-structured system, and the operation of seemingly different networks has universal rules. The wide-ranging research following Barabási (2002) and colleagues (e.g. Albert et al., 1999; Jeong et al., 2000) - innovative work has pointed out that cells, living organisms, the Internet, economy, etc. all make up networks, and they all work according to the same general laws. Barabási referred to these scientific results in his book, published in several languages. The networks of business and economy are amazingly numerous. Networks can be found in politics; there are property relation networks, co-operation networks, institution networks, network marketing - anything one can imagine. It would be impossible to unite all these interrelationships in a single, comprehensive network. Yet the same rules everywhere regulate natural networks, regardless of the observed structural level. The challenge both for economic and network research is to put these laws into practice.

Among networks, scale-free networks have been especially under the lens, as these systems are very frequent in our world.

Barabási (2002) writes about the operation of the economy, that, in fact, a market is nothing more than a directed network. Companies, groups of companies, financial institutions, governments and each possible participant of the economy are nodes in this network. The links show the relations among these institutions, including purchases and sales, shared research and marketing programs, etc. The importance of a relationship shows the value of the transaction between two units, and these connections also possess a direction: namely from the sender to the receiver. The structure of this directed and weighted network determines the outcome of macro-economical processes. A similar statement can be made about almost all areas of our environment, which demonstrates the general need for researching and teaching network science.

[*]University of West Hungary, TermIK; Hungary
[†]University of Pécs, FEEK; Hungary
[‡]University of Pécs, FEEK; Hungary
[§]University of Pécs, FEEK; Hungary

Following, we will examine recent problems in the acquisition of knowledge concerning development and networks. We mean by the term *acquisition of knowledge* the process mentioned above, by which individuals can access what is required from society's store of knowledge. At this point we only draw attention to the fact, that in the near future fruitful research in social and human science will have a much bigger significance than in former periods. This statement is especially relevant for interplay between language and intercultural communication, chiefly for researches of foreign languages acquisition.

2 Network Elements

If we look at a map, we can see a road-network in which the nodes are the settlements, and the edges are the roads connecting them. We can build up a similar, so-called acquaintance-network if we put certain people's names in the nodes, and the edges are drawn between those who know each other. Networks can also be drawn to show the connections between a company and other institutions, or many other relationships.

A general interpretation of networks can be given in the following way: the nodes of a network represent the elements of a set, and the edges (or links) represent a certain relation connecting these elements. In the former examples, settlements, people, and institutions make up the set, and the relations between them are the traffic possibilities (i. e. roads), acquaintanceships, and official relationships. Our concepts of encoding knowledge and our terms for ensuring linguistic communication also make networks. The edges of a network can mark both reciprocal and directed relationships.

Network theory was originally elaborated in mathematics, more specifically in the field of graph theory. The model for the research was created by making random connections among elements by throwing a dice according to certain rules. All the nodes in these random networks have the same characteristics. The existence of a network-structure (that is, graphs) among the elements of certain sets was discovered long ago in several areas. However, it was only a decade ago that researchers noticed the significance of network-like operation and the limits of the random network model. The foundations of the modern theory of random graphs were laid in the 1960s in eight studies published by Rényi and Erdős's. Karonski and Rucinski (1997) provide an evaluation and enumeration of these foundations in their study.

Attention was drawn to the special importance of networks by a game programme created by American university students'. The essence of this game was to find the shortest possible connection between given Hollywood movie stars through a series of actors they starred together with in Hollywood films. The surprising result of the game was that the problem could be solved through on average six connected pairs in every case. The fact that there exists a short series of connections between two seemingly far points of the network suggests that our knowledge about interaction and connection of different things have a completely different nature than we used to believe. Research results show that in every network, and thus in every natural or social system, *six degrees of separation* can be found, that is, they make up an easily navigable *small world*. Science has served the purpose of recognising parts of reality and details of the network so far. The task in the near future will be to put together these details to gain knowledge about the operation of the whole network.

The recognition of the role of networks, i.e. the fact that everything includes a network structure, brought research on network characteristics into the spotlight. The World Wide Web is a network in which websites are the nodes, and possible links are the edges. This network functions in a digital system, thus it provides an excellent opportunity for being used as a model system. A part of this network has been mapped, and the first analyses showed that this functioning network differs in many aspects from the one described by the random model mentioned earlier. For example, concerning the distribution of the converging edges, we would expect the number of nodes with 1, 2, 3, 4, etc. links to increase for a while, then to decrease after reaching the maximum. In reality, the number of links of websites does not show a maximum value - i.e. their distribution is characterised not by a *bell curve*, but by an abruptly rising *curve of a power law*. The other difference between the random network and the Web is that while the nodes of the former are similar to each other - in other words, there is an internal scale by which they can be measured - there are enormous differences between the nodes of the latter. Besides a great number of nodes with few links, there are a few so-called *hubs* with several links. These hubs play a special role in network formation and functioning. The group of complex networks in which the number of different nodes is described by *a power law distribution* and the nodes cannot be categorised by an internal scale are called scale-free networks. Continental flight route maps are typical *scale-free networks*, in which the nodes are the airports and the edges are the flight routes. We can clearly see that the nodes play different roles in air traffic, and the central airports function as hubs (Amaral et al., 2000).

In the following we will touch upon some basic features essential to the application of scale-free networks (for details e.g. Barabási, 2002; Csermely, 2005; Kogut & Walker 2001).

- The characteristics of expansion rules of networks. These networks build up node by node, in a self-organising way. During development all networks, either biological systems, social relationships, economical or regional networks, started from a small core, and the network in built up by adding new nodes. A new node is likely to connect to a network node with several existing connections. This is called the *principle of popularity*, and is realised during expansion. This law manifests itself in the centralised growth of today's economic networks in the predominance of the principle "the rich gets richer".

- Another important question is the networks' fault tolerance. The question is: if one or more nodes drop out of operation, how long will the remaining network continue working? Natural systems can generally continue to work after an element drops out, while synthetic constructions become inoperable at the failure of an element. Examination of random networks has shown that the elimination of some nodes hardly affects the functioning of the network. However, if the number of faults reaches a critical value, the network disintegrates. The Internet-map model experiment examining scale-free networks led to a surprising result: with eighty percent of the nodes removed, the remaining part was still functioning. Further model experiments showed the outstanding role of hubs in the fault tolerance of the system. However, the failure of a relatively small number of hubs led to the disruption of the network. The predominant role of hubs in the vulnerability of scale-free networks is a typical feature of these systems.

- Not only the nodes of scale-free networks, but also the edges (i.e., the links between them) differ from each other. We can distinguish between *strong links* and *weak links*. The former are close connections which keep the network together, while the latter are distant connections and ensure the stable operation of the network. Concerning the functioning of the network both types are equally important.

- Due to their versatile features, every element of the entities composing our environment can be a node in several networks at the same time. This means different networks are interconnected. Therefore, if we wish to illustrate the structure and functioning of language networks, a complicated spatial model needs to be created. (see Fóris 2007b).

- The set of elements containing nodes determines the type of the network, which can be spatial, temporal, conceptual, etc.

The recognition that network-like functioning is present in natural, social, and economical processes has, together with earlier mathematical research, promoted searching for applications based on the laws of scale-free networks.

3 The role of the networks in the knowledge transfer process

The knowledge possessed by people living in social communities has always had two components. One part of this knowledge can be conceived verbally or in written form, or recorded in the form of images; thus it can be stored, transported, and assigned to be learned. We call the attainments recorded this way *encoded knowledge*. Its transfer devices are books, patents, technical descriptions, films, musical scores, tapes, etc. This part of our knowledge can be taught as lessons to be learned, it is recognisable, and its acquisition can be quite precisely checked by testing and examining. The other part of our knowledge cannot be encoded this way - it emerges when we use our attainments It is hard to detect separately; it cannot be recorded, transported, or taught as a separate task. This part of our knowledge is called *tacit knowledge*, and its constituents are *abilities* (problem solving, decision making, estimation of situations, etc.), *aptitudes* (diligence, creativity, work capacity, manners, etc.), *intelligence*, and *erudition*. It is important to understand that an individual's tacit knowledge depends partly on his/her unique abilities, and partly on acquired life experiences. Individuals can be helped in this acquisition. Practice, sharing experience, and training all serve this purpose, but it cannot be taught directly as encoded knowledge. The use-value of knowledge is given by the simultaneous presence of these two knowledge types as relevant to the given task. During our lives we need to use tacit and encoded knowledge together.

At the early stage of development these two types of knowledge and their accessibility were not clearly separated. For example, members of nomadic (and later agricultural) societies possessed vast amounts of knowledge about the biology and physiology of plants and animals, architecture, human and animal health and healing, different ways of transferring knowledge (and the importance of doing so), and many other things. These acquirements were transferred within the members of the community. In modern words, it was real practice-oriented education. Formal education had not been organised yet. The knowledge gathered by the community was not accumulated in a concentrated way, but in a self-organising network in which the members of the community were the nodes.

In this network knowledge was available for everyone; everyone could access it depending on his/her needs in life. Knowledge-transfer did not have formal, institutional forms; it took place practically through the community network. The nodes of the network represented specific individuals' store of learning, while edges represented the possible connections. An individual's knowledge slightly differed from that of the others, but the knowledge stored in different places was widely accessible. The network stored all the experience the community needed. If some knowledge needed to be complemented at any point of the network, knowledge transfer could happen quickly and simply. Thus, at this stage of development, the spatial distribution of knowledge was homogenous. The difference in profundity was not significant. At this point, professional differentiation could be found only in traces.

With the development of social labour division, and the increasing amount of knowledge and its shifting depth, acquisition had to be segmented. Crafts, more and more independent medicine, controlling warcraft on different levels, developing government, etc. all required unique training. The encodable part of knowledge became significant; above all writing, reading, calculating and a part of professional attainments belonged to this group. There also was a need for encoding and storing knowledge; the development of writing and later printing were great leaps forward. The system of books and libraries evolved, and they became a source of the knowledge accumulated by the society. An organised system of transferring encodable knowledge also was forming gradually, and instructors possessing a vast amount of knowledge gained a special role in it. The homogeneous distribution of the knowledge accumulated by society was replaced by differentiation based on social levels, and knowledge was concentrated at certain points. The way of attaining knowledge also changed. Individuals could access only the part of society's knowledge they required for their individual goals.

School network and vocational training system developed during this scientific and professional differentiation process transferred the bulk of encoded knowledge to students (Sári, 2004). Nevertheless, non-formal channels of attaining at significant amount of tacit knowledge also persisted. Through these people acquired their first language, family and social relations, and social and moral norms. Later these channels also took over transferring new attainments, which were being produced more and more quickly (Koltai, 2002; Kozma, 2002).

Due to accelerating scientific and technical development of the 20th century, in the last third of the century several important changes emerged in the ways of attaining knowledge, and in the conditions and distribution of application. These were the following: (1) an increase of the amount and standard of encoded knowledge, (2) a similarly rapid growth in tacit knowledge, (3) transfer channels emerged outside of formal education for transferring encoded knowledge, (4) the rate of re-education and continuative education following basic training increased, and networks were developed for these purposes, (5) professional integration clearly emerged in professional-social activities, which required the possession of knowledge from several professions to a certain degree, (6) a series of new professions appeared; in most cases on the borders of traditional fields, (7) traversability among different fields required the diffusion of rigid channels of the knowledge-transfer system, (8) the parallel-linear structure of education was not able to answer the demands of the modified conditions.

The situation developed by the turn of the millennium and described in the above eight points, sets new requirements for the organisation of knowledge-distribution and for the possibilities of its access. Due to professional integration, significantly more profound general and professional attainments are required in larger quantities almost all fields. The set of comprehensive knowledge has been extended, with areas such as information technology, foreign languages, questions of production and government, etc. The possession of tacit knowledge at a much higher level is necessary in all areas. The requirements of abilities and expertise are of a different quality, and these determine the criteria of qualification, especially for top positions. However, development not only requires a higher level of knowledge, it also provides opportunities for its access. Digital technology, which is generally applied for recording, storing, and transferring information, is radically altering the ways of attaining knowledge.

As we have seen, the functioning of developed societies is characterised by increasing professional integration and its consequence is that any specialist needs more and more skills other than the attainments of his or her own profession. For example, information technology, native and foreign languages are required besides professional skills in every vocation. Technical, commercial, legal, and economic attainments and proficiency are required for the technical estimation and acquisition of devices, and for being aware of the legal and economical terms and conditions of functioning. The means of transferring this knowledge is language (in the case of international communication it is the foreign languages), and we have to learn about the network structure of language approximately convey information.

This situation, demanding integrated knowledge, is very similar to the initial phase of development. Individuals or groups working on a certain task have to possess complementary attainments from different fields. The want of new attainments occurring during work needs to be solved in a short time. This demand can be fulfilled only if the knowledge accessible for society is stored in a network-like distribution and differentiated on the scientific level, which ensures quick access for its users. The linear knowledge storage and transfer does not ensure the possibility of the necessary selection from the stored knowledge. In the linear system every intermediary level presumes the detailed knowledge of its precedents. If users with another profession wish to acquaint themselves with a detail they need or with the essence of the attainments, they have either to acquire the details of the preliminary knowledge, or abandon the acquisition of the questions which are important for them. This difficulty can be solved if the available knowledge is stored in a modular system. Modules are segmented depending on the users' knowledge level, and the volume and storage method of the knowledge.

Before networks were widely researched, there were a great number of linguistic findings on natural languages that now lend themselves to new interpretation through network theory. Now that network theory has been widely interpreted and mathematically described, previous linguistic findings can be inserted into the system of networks (for details see Fóris, 2007a). The mathematical calculations support the hypothesis that language networks can be described with the help of scale-free networks. It is also evident that the power law of distribution over the network not only describes the frequency of the use of words, but the distribution of other properties as well, such as the distribution of the number of meanings to a word, the parts of speech and the proper nouns in a corpus (Fóris, 2007b).

The exact demonstration of the conceptual system of knowledge processing, the introduction to the concepts and the terminological system based on scientific attainments, the consistent encoding of attainments, the operation of the vehicles of knowledge play an important role in the establishment of skills for attaining knowledge stored in the knowledge network; the modern theories of *terminology* and *ontology* deal with these question. Further exploration of this aspect may be found in many sources (e.g. Cabré Castellví, 2003; Chandrasekan et al., 1999; Miller et al., 1990; Prószéky & Miháltz, 2002; Temmerman 2000).

The discovery of the scale-free nature of language networks demands a new approach in linguistic research. It is always observable in nature, the economy, microscopic and cosmic systems that everything is connected with everything else. That is why, in conjunction with localised research, a special emphasis is currently laid on the approach to focus on the interrelations of complex systems. This leads to the compilation of modern computerised ontology, as beyond providing factual knowledge, it is also practical to provide the environmental links of that knowledge. Within the complex system of language networks, sub-networks are also interrelated, and therefore it is necessary to study the relations of the various units.

4 Summary

In this study I briefly touched upon the main characteristics of the scale-free network model that can be widely applied in knowledge transfer process by language networks. Knowledge is made accessible by storing it in networks. Attainments considered adequate for users' demands have to be stored in the nodes of the network, so that everyone can access them through the network.

References

[1] Albert, R., Jeong, H. and Barabási, A-L. (1999) Diameter of the World Wide Web. *Nature* 401, 130-131.

[2] Amaral, L. A. N., Scala, A., Barthélémy, M. and Stanley, H. E. (2000) *Classes Of Small-World Networks*. Proceedings of the National Academy of Sciences. 97, 11149-11152.

[3] Barabási, A-L. (2002) *The New Science of Networks*. Cambridge, MA.: Perseus.

[4] Cabré Castellví, M. T. (2003) Theories of Terminology. Their description, prescription and explanation. *Terminology* 9 (2), 163-200.

[5] Chandrasekan, B., Josephson, J. R. and Benjamins, V. R. (1999) What are ontologies, and why do we need them? *Intelligent systems* 14 (1), 20-26.

[6] Csermely, P. (2005) *A rejtett hálózatok ereje*. Budapest: Vince Kiadó.

[7] Fóris, Á. (2007a) A skálafüggetlen hálók nyelvészeti vonatkozásai [Linguistic Applications of Scale-free Networks]. *Alkalmazott Nyelvtudomány* VII (1-2),

[8] Fóris, Á. (2007b) Terminology and the Theory of Scale-free Networks. In: Á. Fóris and J. Pusztay (eds.) *Proceedings of the international conference "Current Trends in Terminology"* (Terminologia et Corpora IV) (in press). Szombathely: BDF.

[9] Jeong, H., Tombor, B., Albert, R. Oltvai, Z. N. and Barabási, A-L. (2000) The Large-Scale Organization of Metabolic Networks. *Nature* 407, 651-654.

[10] Karonski, M. and Rucinski, A. (1997) The Origins of the Theory of Random Graphs. In: R.L. Graham and J. Nesetril (eds.) *The Mathematics of Paul Erdős* (pp. 331-336). Berlin: Springer.

[11] Kogut, B. and Walker, G. (2001) The Small World of Germany and the Durability of National Networks. *American Sociological Review* 66, 317-335.

[12] Koltai, D. (2002) *Theoretical, economic and regional issues of adult education. Hungarian developments in an international perspective*. (International perspectives in adult education 33) Bonn: IIZ/DVV.

[13] Kozma, L. (2002) *Tudomány - Technika - Társadalom*. [Science - Technique - Society] Pécs: PTE FEEFI.

[14] Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., and Miller, K. J. (1990) Introduction to WordNet: an On-line Lexical Database. *International Journal of Lexicography* 3, 235-244.

[15] Prószéky G. and Miháltz M. (2002) Automatism and user interaction: Building a Hungarian WoldNet. In: *Proceedings of the 3rd International Conference On Language Resources and Evaluation*. (LREC 2002) Vol. III. (pp. 957-961). Las Palmas.

[16] Sári, M. (2004) *A kultúra intézményrendszereinek történeti-funkcionális változásai*. Pécs: PTE FEEFI.

[17] Temmerman, R. (2000) *Towards New Ways of Terminology Description. The sociocognitive approach*. (Terminology and Lexicography Research and Practice 3) Amsterdam/New York: John Benjamins.