The global organization of the English lexicon and its evolution

Mieko Ogura & William S-Y. Wang
Tsurumi University, Yokohama / Chinese University of Hong Kong*

We present a quantitative study of the semantic network of the set of nouns and verbs of WordNet, which is a systematic representation of the Present-day English lexicon based on psycholinguistic considerations, and *A Thesaurus of Old English* to understand the evolution of the global organization of the English lexicon. We demonstrate that whereas the semantic network is dominated by the hypernymy tree, which works as the skeleton of the set of nouns and verbs, the inclusion of polysemy produces a drastic global reorganization of the semantic structure, that is, it is converted into a small world, where all meanings are closer to each other. We then show that the words with higher frequency and therefore with higher number of meanings construct the higher level of the hypernymy tree within each lexical category. This architecture is robust through the times, forming the basis of the small-world network. We also suggest that the small-world topology of the brain has enhanced the small-world configuration of semantic structure.[1]

1. Theoretical preliminaries

The general properties of the organization of social and biochemical networks have been characterized with the graph theoretic measurements, revealing common features of self-organized systems of highly connected elements.[2] Ogura & Wang (2008b) examine how different structures of social networks affect

---

* Mieko Ogura & William S-Y. Wang are also affiliated to the Project on Linguistic Analysis, University of California at Berkeley.

1. This work is supported by the grants from the Human Frontier Science Program and the Ministry of Education, Culture, Sports. Science and Technology of Japan. We wish to thank Michiyo Takeda for the data collection and Minoru Ichijyu for the computer programming. We are also grateful to the editors for their helpful comments.

2. Watts & Strogatz 1998, Watts 1999, Barabási & Albert 1999.

linguistic selection type of change and game type of change, based on simulation and historical data from English.

In this study we use similar tools to show several global properties in the evolution of the English lexicon, based on the set of nouns and verbs of WordNet (version 2.0) and *A Thesaurus of Old English* (*TOE*) (1995). We further Sigman and Cecchi (2002) and demonstrate that whereas the semantic network is dominated by the hypernymy tree, which works as the skeleton of the set of nouns and verbs, the inclusion of polysemy produces a drastic global reorganization of the semantic structure, that is, it is converted into a small world, where all meanings are closer to each other. We then examine the effects of word frequency on the small-world semantic network. We also suggest the implication of small-world networks of the lexicon for brain networks.

WordNet is a systematic representation of the English lexicon based on psycholinguistic considerations. Its design resembles a thesaurus in that its building block is a set of synonyms. The sets of synonyms are organized as a lexical hierarchy of hypernyms (Fellbaum ed. 1998). *TOE* uses as its main source material the word senses from the *Oxford English Dictionary* (*OED*) and standard Anglo-Saxon dictionaries. The overall structure of the classification is hierarchical, proceeding from the most general terms to the most specific.

A graph, in its most basic sense, is a set of points connected by a set of lines. Graphs can be used to represent all kinds of networks (Watts 1999). Based on the analysis of large-scale complex networks in the real world, Watts & Strogatz (1998) and Watts (1999) proposed a model of a graph for a small-world network that reconciled the high clustering of regular network with the haphazard character of random network. They started from a circle of nodes, where each node is connected to its immediate and next-nearest neighbors as shown in a regular network in Figure 1A below. To make this world a small one, a few regular links were rewired, connected to randomly selected nodes as shown in a small-world network in Figure 1A. These long-range links offer the crucial shortcuts between distant nodes, shortening the average length between all nodes. Watts (1999) presents a lattice substrate as shown in Figure 1A and a tree substrate as shown in Figure 1B below for the models of graphs of the small-world networks. A regular network in Figure 1B shows a tree substrate with branching of 2. To make this world a small one, nodes are connected randomly by the links as shown in the solid lines in a small-world network in Figure 1B.

Dictionaries make implicit use of the hypernym relationship in defining a word by its hypernym and its specific attributes. Thus we use the hypernymy tree as a base graph. In Figure 1B broken lines represent the hypernymy tree. The lexicon then defines a graph, where the nodes are the semantic categories or meanings composed of a set of vertices, i.e. synonyms. Semantic relationships are the

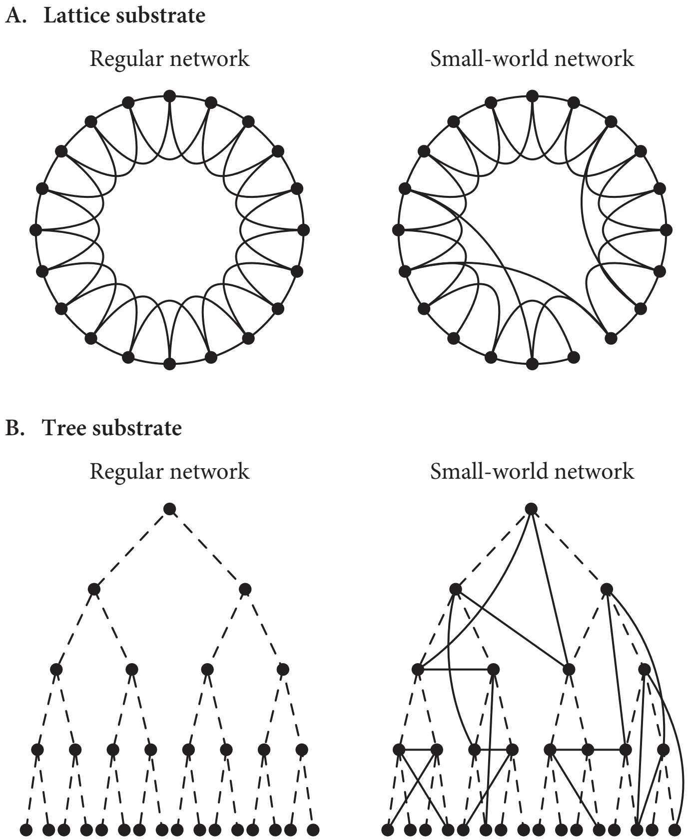


Figure 1. Two types of networks with a lattice substrate and a tree substrate (Adapted from Watts 1999)

links. In Figure 1B the solid lines represent the polysemous links. Graph theory provides a number of measurements that characterize the structure of a graph: the characteristic length, which is the median of the distribution of average minimal lengths across all vertices, the distribution of links, i.e. first neighbor connections, and the clusters, which define regions of very high internal connectivity.

The methods of computation in the present study are as follows. We assume that if the lexicon is composed of monosemous words alone, the length from word i to the synonymous words within the semantic category to which i belongs is 0. The distance from i to the synonymous words in different semantic categories is calculated by climbing up and down the hypernymy tree. In this way the semantic categories connected by the hypernymy tree form vertical networks.

When polysemous words are included in the lexicon, the length from word i to the synonymous words within the semantic category A to which i belongs is

0, and that from i to the semantic category B to which the semantic category A is linked through a polysemous word j is 1. If word k in the semantic category B is a polysemous word which is connected to the semantic category C, the distance from i to the synonymous words in the semantic category C is 2. This process continues until there is no polysemous word in the semantic category. In this way semantic categories connected through polysemous words form horizontal networks. For a polysemous word and a monosemous word that joins horizontal networks through polysemous words, the distance via the horizontal networks and that via the vertical networks are compared and the shorter value is adopted. For other monosemous words the distance is calculated via vertical networks.

For a word, the average minimal length is calculated by averaging the minimal distance from a word to all of the other words. The characteristic length is computed as the median of the distribution of average minimal lengths across all words.

As for links, we assume that word i is linked to the synonymous words in the semantic category (categories) to which i belongs, and to the synonymous words in the hypernym immediately above the semantic category (categories) to which i belongs. As for clusters, we assume that if semantic categories are connected through a polysemous word, synonymous words in the semantic categories to which a polysemous word belongs form the maximal possible number of connected neighbors. If the semantic category is not connected to the semantic categories, synonymous words in that semantic category form the cluster.

We consider the global organization of the English lexicon from an evolutionary perspective. Hurford (2007) considers that the ability to form complex conceptual structures is crucial to the emergence of human language. He asserts that no such complex communication system could have evolved without reliable cooperativeness, and suggests Tomasello et al. (2005)'s concept of shared intentionality as a key ingredient of humans' striking willingness to play complex language games with each other.

The crucial last biological step towards modern human language capacity was the development of a brain capable of acquiring a much more complex mapping between signals and conceptual representations, giving rise to the possibility of the signals and the conceptual representations themselves growing in complexity. In keeping with ideas from grammaticalization theory about meaning, the earliest languages would have had, in their semantics: no metaphors; no polysemy; no abstract nouns; fewer subjective meanings; less lexical differentiation; fewer hyponyms and superordinate terms (Hurford 2003).

Within this evolutionary perspective, we investigate how lexicons have organized through Old English to Present-day English. *TOE*, our database for

Old English uses as its main source material the word senses from the *OED* and standard Anglo-Saxon dictionaries. These dictionaries are based on some 2,000 surviving Old English texts. In spite of the limitations of Old English data, it is clear that the numbers of words and meanings have increased from Old English to the Present-day English. The lexicon itself has grown in complexity. In this study we present a quantitative study of the graph structure of the set of nouns and verbs in WordNet and *TOE* to understand the evolution of the global organization of the English lexicon.

2. Small-world networks of nouns and verbs in WordNet

We have analyzed 11,306 (6,110 monosemous + 5,196 polysemous) verbs in WordNet. The average number of meanings of polysemous verbs is 3.56. Figure 2 shows that the inclusion of polysemy results in a small-world organization of the semantic graph, where all meanings are closer to each other.

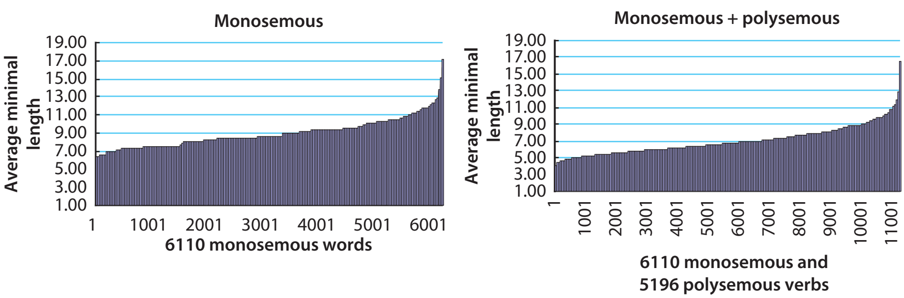


Figure 2A. The histogram of average minimal length of the monosemous (left) and monosemous + polysemous (right) verbs in WordNet

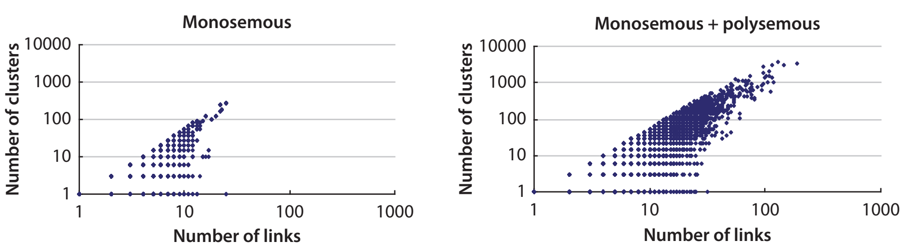


Figure 2B. The plot of the clusters as a function of links of the monosemous (left) and monosemous + polysemous (right) verbs in WordNet

Figure 2A is the histogram of average minimal length arranged from the shortest and Figure 2B is the plot of the connected neighbors, i.e. clusters as a function of neighbors, i.e. links for 6,110 monosemous verbs (left) and 6,110 monosemous + 5,196 polysemous verbs (right). 6,466 out of 11,306 verbs form horizontal networks. The average number of verbs whose distances are calculated via horizontal networks for the average minimal length of each verb is 2,776. We can see that the inclusion of polysemy decreases the average minimal length and produces a considerable number of verbs with high connectivity and clustering.

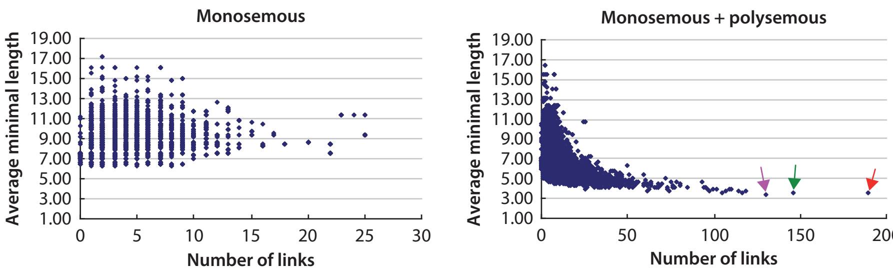


Figure 2C. The average minimal length as a function of links of monosemous (left) and monosemous + polysemous (right) verbs in WordNet

Figure 2C shows the average minimal length as a function of the number of links for monosemous verbs and monosemous + polysemous verbs. The figure for the monosemous + polysemous verbs shows power-law distributions of both average minimal length and the number of links. The verbs with the greatest number of links (L) form hubs (marked with arrows), which correspond to the verbs *break*, *make*, and *get* (from the left). These three verbs are the most polysemous (P) and frequent (F) (for word frequency, see Section 4) words with the shortest average minimal length (M), respectively: *break* (L = 189, P = 59, F = 99, M = 3.6), *make* (L = 146, P = 49, F = 1638, M = 3.59), *get* (L = 130, P = 36, F = 732, M = 3.4).

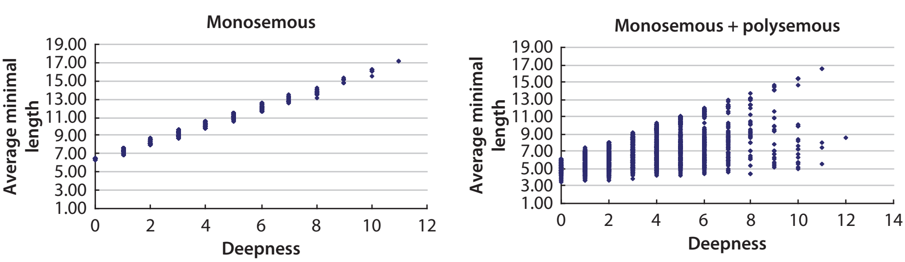


Figure 2D. Correlation between average mean length and deepness in the hypernymy tree of monosemous (left) and monosemous + polysemous (right) verbs in WordNet

Figure 2D displays the correlation between average mean distance and deepness in the hypernymy tree with and without the inclusion of polysemy. In monosemous verbs there is a correlation between average mean distance and deepness. When polysemy is added, the correlation, though still present, becomes very weak, showing the hierarchy of hypernymy has low impact on distance of meanings.

To summarize, the inclusion of polysemy changes average minimal length, number of neighbors and connected neighbors and has a profound impact on the organization of the lexicon. The most important change in characteristic length results from adding polysemy. In addition, the inclusion of polysemy produces a considerable number of words with high connectivity and clustering. On average, the inclusion of polysemy reduces the characteristic length from 8.92 to 6.98, and increases the number of neighbors and connected neighbors from 4.2 to 8.2, and from 8.3 to 27.53 respectively. The inclusion of polysemy creates a clustered short-range, i.e. small-world semantic network.

We have also analyzed 114,513 (99,234 monosemous + 15,279 polysemous) nouns in WordNet (Version 2).[3] The average number of meaning of polysemous nouns is 2.77.

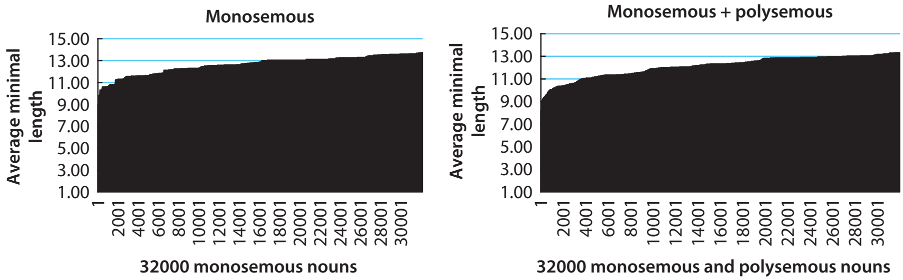
Figure 3 shows (A) the histogram of average minimal length arranged from the shortest for the first 32,000 monosemous words and monosemous + polysemous words arranged alphabetically, (B) the plot of the clusters as a function of links, (C) the average minimal length as a function of the number of links, and (D) correlation between average mean distance and deepness in the hypernymy for the first 60,000 monosemous words and monosemous + polysemous words arranged alphabetically. The nouns with the greatest number of links form hubs, which correspond to the nouns: *line* and *head* (from the left). These 2 nouns are the most polysemous and frequent words with the shortest average minimal length, respectively: *line* (L = 119, P = 29, F = 192, M = 7.77), *head* (L = 94, P = 32, F = 254, M = 7.10). It is to be noted that *head* forms a hub from OE (see Section 3) through Present-day English.[4]

The inclusion of polysemy creates a small-world network of nouns, but the degree of small world is less in nouns than verbs in WordNet. On average, the inclusion of polysemy reduces the characteristic length from 13.96 to 13.46, and increases the number of links and clusters from 2.27 to 4.19, and from 2.07 to 5.08 respectively. 11,700 out of 114,513 nouns form horizontal networks, and the average number of nouns whose distances are calculated via horizontal networks

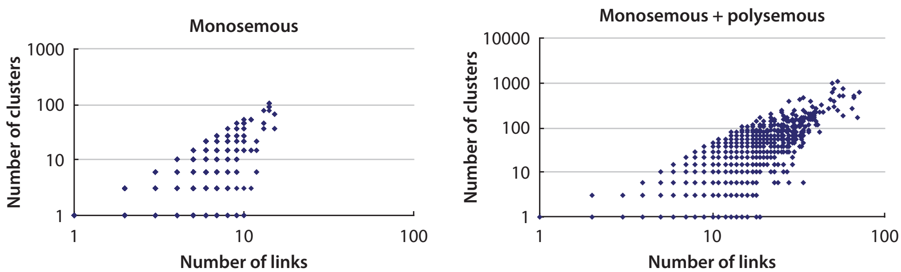
3. We have deleted 135 numerals from our database.

4. See also Healey 2011.

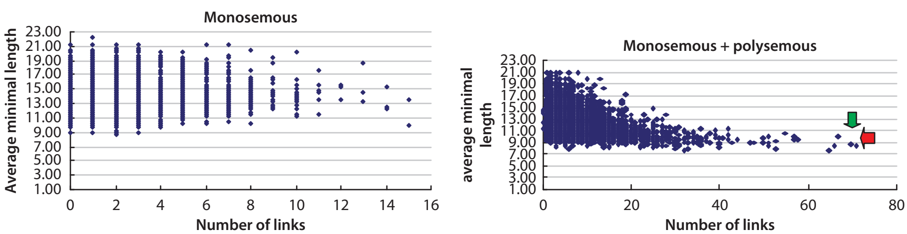
(A) The histogram of average minimal length of monosemous (left) and monosemous + polysemous (right) nouns



(B) The plot of clusters as a function of links of monosemous (left) and monosemous + polysemous (right) nouns



(C) The average minimal length as a function of the links of monosemous (left) and monosemous + polysemous (right) nouns



(D) Correlation between average mean distance and deepness in hypemymy of monosemous (left) and monosemous + polysemous (right) nouns

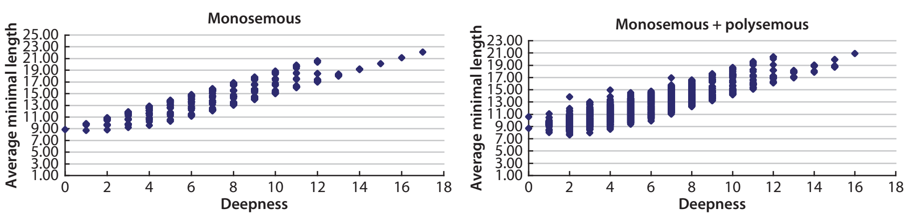


Figure 3. Small-world networks of nouns by the inclusion of polysemy in WordNet

for the average minimal length of each noun is only 908. The average number of synonyms in a given semantic category is almost the same in nouns and verbs: 1.44 (max. 28) in nouns and 1.82 (max. 24) in verbs. Thus the less degree of small world in nouns than verbs is due to the low percentage of the polysemous words in nouns.

Sigman & Cecchi (2002) state, based on the 66,025 nouns in WordNet (version 1.6), that the inclusion of polysemy reduces the characteristic length from 11.9 to 7.4. The difference of the results is probably due to Sigman & Cecchi's database and method of computation. The database for version 2.0 is nearly twice as large as that for version 1.6. We assume that the data added in version 2.0 are more specific monosemous words that entered in more recent times (see Section 4). Sigman & Cecchi's minimal distance between vertices was computed adapting a publicly available version of Dijkstra's algorithm.

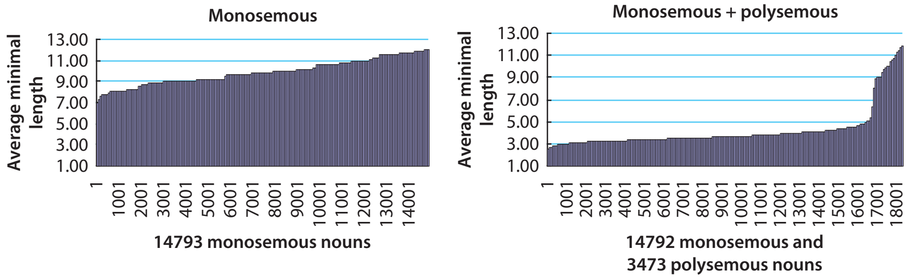
3. Small-world networks of nouns and verbs in TOE

In *TOE* the defining headings of the word senses, which are based on wording of the Clark Hall and Bosworth-Toller dictionaries, are written so as to match grammatically what is defined. Thus the wording of the definition reflects the part of speech: verb definitions operating with 'To […]' and adjective ones with adjectival forms. The sense definition for a group of nouns may point to their use in mass or count contexts. First we computerized the data for all the nouns and verbs in *TOE* based on the defining headings of the word senses.

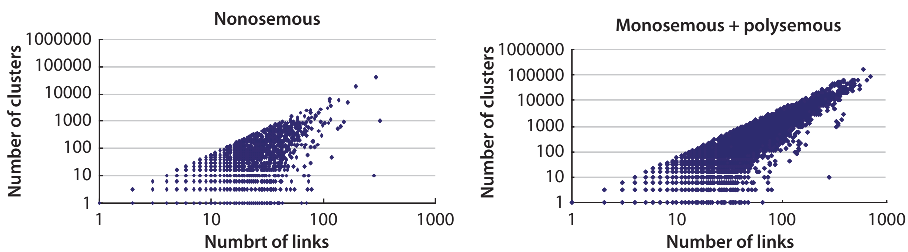
Then we analyzed all the nouns and verbs in *TOE*: 18,265 (14,792 monosemous + 3,473 polysemous) nouns and 7,161 (5,019 monosemous + 2,142 polysemous) verbs. The average numbers of meanings of polysemous nouns and verbs are 2.9 and 3.14 respectively. 16,765 out of 18,265 nouns and 6,825 out of 7,161 verbs form horizontal networks. The average numbers of nouns and verbs whose distances are calculated via horizontal networks for the average minimal length of each noun and verb are 16,477 and 6,660 respectively. In *TOE* the semantic categories are represented by the numerical hierarchy and are given if at least one of the nouns, adjectives, adverbs, verbs and others occurs as a hypernym. Thus in our calculation of the minimal distance via vertical networks, if the nouns or verbs do not occur in certain semantic categories in the hierarchy, we do not count them as hypernyms.

Figures 4 and 5 show (A) the histogram of average minimal length arranged from the shortest, (B) the plot of the connected neighbors, i.e. clusters as a function of neighbors, i.e. links, (C) the average minimal length as a function of the number of links, and (D) correlation between average mean distance and deepness in the hypernymy for monosemous words and monosemous + polysemous words in

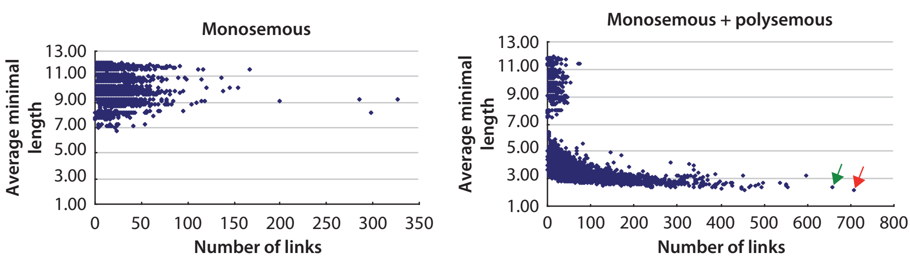
(A) The histogram of average minimal length of monosemous (left) and monosemous + polysemous (right) nouns



(B) The plot of clusters as a function of links of monosemous (left) and monosemous + polysemous (right) nouns



(C) The average minimal length as a function of links of monosemous (left) and monosemous + polysemous (right) nouns



(D) Correlation between average mean distance and deepness of monosemous (left) and monosemous + polysemous (right) nouns

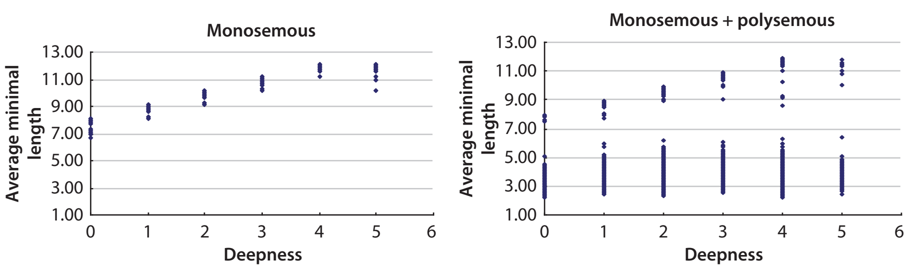
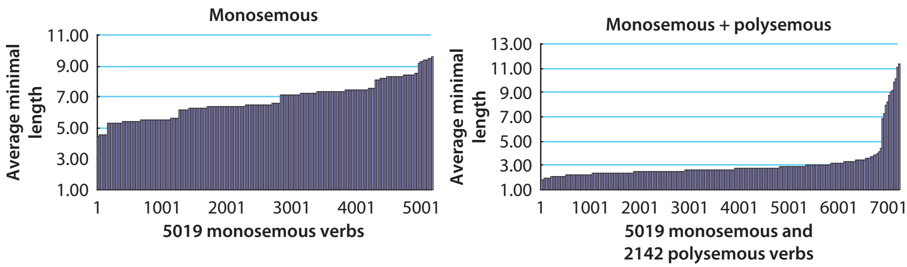
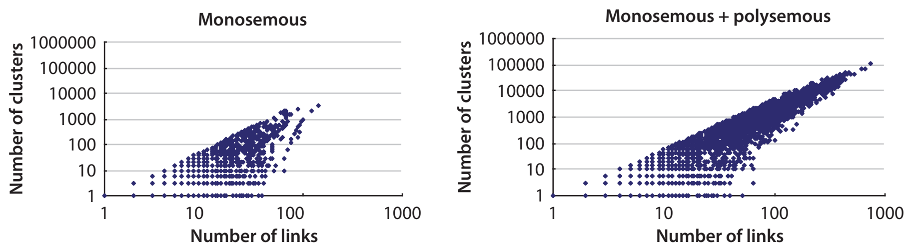


Figure 4. Small-world networks of nouns by the inclusion of polysemy in *TOE*

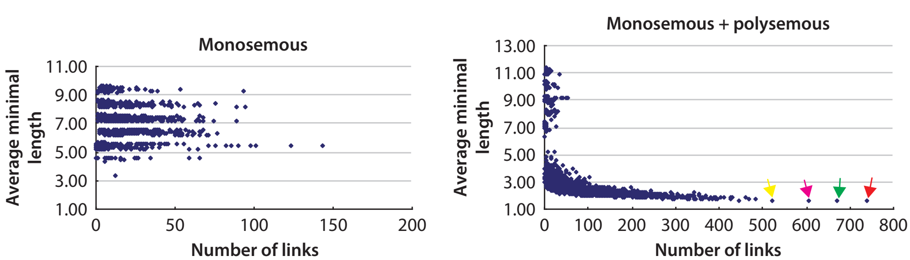
(A) The histogram of average minimal length of monosemous (left) and monosemous + polysemous (right) verbs



(B) The plot of clusters as a function of links of monosemous (left) and monosemous + polysemous (right) verbs



(C) The average minimal length as a function of links of monosemous (left) and monosemous + polysemous (right) verbs



(D) Correlation between average mean distance and deepness in the hypemymy monosemous (left) and monosemous + polysemous (right) verbs

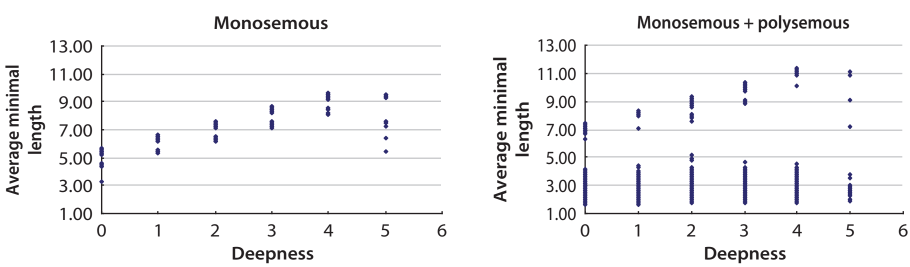


Figure 5. Small-world networks of verbs by the inclusion of polysemy in *TOE*

nouns and verbs respectively. In Figure 4C the nouns with the greatest number of links form hubs (marked with arrows), which correspond to the nouns: *hlāford* 'lord', *hēafod* 'head' (from the left). These 2 nouns are the most polysemous and frequent words with the shortest average minimal length, respectively: *hlāford* (L = 657, P = 14, F = 632, M = 2.41), *hēafod* (L = 707, P = 17, F = 814, M = 2.17).[5] In Figure 5C the verbs with the greatest number of links form hubs (marked with arrows), which correspond to the verbs: *(ge)healdan* 'hold', *(ge)niman* 'take', *began* 'bego' and *āwendan* 'awend' (from the left). These 4 verbs are the most polysemous and frequent words with the shortest average minimal length, respectively: *(ge)healdan* (L = 522, P = 21, F = 742, M = 1.62), *(ge)niman* (L = 604, P = 21, F = 230, M = 1.63), *began* (L = 670, P = 15, F = 139, M = 1.68), *āwendan* (L = 740, P = 17, F = 94, M = 1.6).

The inclusion of polysemy results in a clustered and compact graph, a small world. The addition of polysemy reduces the characteristic length considerably, but with a considerable increase of links and clustering. On average, the inclusion of polysemy reduces the characteristic length from 9.88 to 4.2, and increases the number of neighbors and connected neighbors from 43.31 to 52.94, and from 1,301.52 to 1,549.66 respectively in the nouns; and it reduces the characteristic length from 6.77 to 2.99, and increases the number of neighbors and connected neighbors from 32.33 to 52.37, and from 316.59 to 11,09.45 respectively in the verbs.

Comparing the results for nouns and verbs in *TOE* with those in WordNet, we find that when polysemous words are included, the decrease in characteristic length and the increase in the average number of neighbors and connected neighbors are greater in *TOE* than WordNet. That is, the degree of small world of the

5. Healey (2011) explores the range of senses (literal, transferred, and figurative) of '*hēafod*', one of the most frequent words in Old English. She shows a large entry of 16 main senses and many sub-senses (Senses 1 through 6 illustrate the literal senses; those 7 though 12 the transferred senses; those 13 through 15 figurative senses; 16 special uses in phrases) in the *Dictionary of Old English* (Cameron, Angus, Ashley Crandell Amos and Antonette diPaolo Healy et al. eds.). Many of its senses are still in use in Present-day English. She also shows a few of the current uses of 'head' which are not in Old English.

Healey states that '*hēafod*' continues vigorously in the language up to the present. In fact, '*hēafod*' propagates a vast network of compounds and related words in Old English, no fewer than 105 in number (105 compounds are listed in the note 19), some of which are still current today: *hēafodece* 'headache', *hēafodland* 'headland', *hēafodlēas* 'headless', *hēafodmann* 'head man', *hēafodwund* 'head-wound', *forehēafod* 'forehead', *behēafdian* 'behead', *behēafdung* 'beheading', among others. We assume that the productivity of compounds with '*hēafod*' is due to the shortest average minimal length, or the closest relation to the other words in meaning of '*hēafod*'.

lexicon is greater in *TOE* than WordNet, though the percentage of polysemous verbs is lower in *TOE* than WordNet. We assume that the interaction between synonymy and polysemy in the horizontal networks is crucial for the degree of small world. The average number of synonyms in a given semantic category is 9.21 (max. 282) for nouns and 8.59 (max. 84) for verbs in *TOE*, and 1.44 (max. 28) for nouns and 1.82 (max. 24) for verbs in WordNet. The higher number of synonyms form the larger horizontal networks.

The characteristic lengths of monosemous nouns and verbs in *TOE* are shorter than those in WordNet, because the older the date of origin of nouns and verbs, the less the number of hypernyms (see Section 4).

4. The effects of word frequency

Tables 1 and 2 show the number of meanings and the word frequencies for the nouns and verbs of WordNet and *TOE* respectively. As for nouns in WordNet, we have analyzed the first 50,000 out of 114,513 nouns arranged alphabetically. As for the word frequencies of WordNet, we sum up the number of occurrence of each meaning of the word given by WordNet, which is based on the number of occurrence in semantic concordances from two textual corpora: 103 passages from the Brown Corpus and Stephen Crane's *The Red Badge of* Courage (1895) (Fellbaum 1998). The frequency counts of the *TOE* words are based on *Old English Corpus on the World-Wide Web* (University of Toronto, 1997). We use the number of occurrence of the infinitives for verbs and nominative singular forms for nouns. We find that the greater the number of meanings the greater the word frequency within each lexical category.

Table 1. Number of meanings and word frequency in nouns and verbs of WordNet

Noun			Verb		
Number of meanings	Number of words	Average word frequency	Number of meanings	Number of words	Average word frequency
1	43714	0.17	1	6110	0.85
2	4016	1.33	2	2508	2.84
3	1227	6.9	3	1094	6.11
4	473	15.06	4	604	15.61
5	227	41.28	5	360	23.41
6–10	308	134.38	6–10	484	56.47
11–32	35	296.83	11–20	111	348.23
			21–59	35	533.29

Table 2. Number of meanings and word frequency in nouns and verbs of *TOE*

Noun			Verb		
Number of meanings	**Number of words**	**Average word frequency**	**Number of meanings**	**Number of words**	**Average word frequency**
1	14792	15.17	1	5019	8.27
2	2158	30.04	2	1131	12.05
3–5	1076	133.01	3–5	810	58.1
6–10	203	337.08	6–10	179	86.9
11–18	36	432.89	11–21	22	335.95

Table 3 classifies the first 50,000 out of 114,513 nouns and 11,306 verbs in WordNet according to the dates of origin that are based on the *Oxford English Dictionary, version 2.0 on CD-ROM (OED2).*[6] It shows the number of words (monosemous words in parentheses), number of meanings, word frequency and number of hypernyms. We find that the number of hypernyms is larger in nouns than verbs, and that the older the date of origin, the more the number of meanings and the word frequency, and the less the number of hypernyms within each lexical category.

Table 3. The effects of word frequency on nouns and verbs in WordNet

Noun				
Origin	**Number of words**	**Number of meanings**	**Word frequency**	**Number of hypernyms**
OE	878(373)	2.62	26.79	5.57
12th c.	133(60)	2.42	20.85	5.38
13th c.	777(332)	2.52	29.86	5.2
14th c.	1699(870)	2.04	7.48	5.41
15th c.	1060(610)	1.85	4.7	5.8
16th c.	2999(2003)	1.56	2.9	5.92
17th c.	3328(2499)	1.41	2.1	6.06
18th c.	2947(2428)	1.25	1.18	6.67
19th c.	10080(9087)	1.08	0.23	6.71
20th c.	7939(7509)	1.06	0.13	6.58
*	18160(17943)	1.01	0.02	6.97

(Continued)

7. 18,160 nouns asterisked in the column of origin for the nouns in WordNet are those that are found neither in the headwords nor in the texts of *OED2*. Most of them are monosemous infrequent words, and we assume that they appeared quite recently.

Table 3. (Continued)

Verb				
Origin	**Number of words**	**Number of meanings**	**Word frequency**	**Number of hypernyms**
OE	671(167)	4.85	102.6	2.16
12th c.	115(36)	3.96	48.1	2.1
13th c.	612(167)	3.64	20.1	2.24
14th c.	1284(421)	2.95	10.6	2.38
15th c.	763(278)	2.58	7.6	2.51
16th c.	1822(830)	2.15	5.1	2.65
17th c.	1495(837)	1.8	1.99	2.57
18th c.	732(449)	1.64	1.27	2.69
19th c.	1886(1269)	1.48	0.87	2.57
20th c.	1144(938)	1.23	0.35	2.72

Table 4 classifies 18,265 nouns and 7,161 verbs in *TOE* according to the date of retention that are based on *OED2*. They show the number of words (monosemous words in parentheses), number of meanings, word frequency and number of hypernyms. We find that the number of hypernyms is larger in nouns than verbs, and that the more frequent the words, the later the date of retention and the more the number of meanings within each lexical category.

Table 4. The effects of word frequency on nouns and verbs in *TOE*

Noun				
Date of retension	**Number of words**	**Number of meanings**	**Word frequency**	**Number of hypernyms**
OE	14643(12635)	1.2	7.2	2.22
12th c.	88(46)	2.15	136.33	1.93
13th c.	223(121)	2.13	58.94	1.96
14th c.	148(68)	2.29	164.34	2.05
15th c.	157(91)	2.03	108.77	2.04
16th c.	118(73)	1.86	36.97	2.31
17th c.	112(85)	1.63	33.29	2.35
18th c.	73(44)	1.93	43.07	2.34
19th c.	1260(834)	1.73	41.84	2.19
20th c.	1443(796)	2.27	194.33	2.29

(Continued)

Table 4. The effects of word frequency on nouns and verbs in *TOE* (Continued)

Verb				
Date of retension	**Number of words**	**Number of meanings**	**Word frequency**	**Number of hypernyms**
OE	4735(3768)	1.33	5.61	1.66
12th c.	60(35)	2.2	8.68	1.65
13th c.	247(126)	2.09	14.43	1.44
14th c.	263(144)	2.13	11.18	1.44
15th c.	157(67)	2.52	27.31	1.33
16th c.	126(61)	2.24	12.81	1.67
17th c.	98(57)	1.98	16.58	1.72
18th c.	45(27)	1.64	7.49	1.73
19th c.	661(377)	2.05	23.36	1.72
20th c.	769(357)	2.55	53.61	1.53

We may state that the words with higher frequency and therefore with higher number of meanings construct the higher level of the hypernymy tree within each lexical category. This architecture is robust through the times, forming the basis of the small-world network. Trees have definite roots and branches that distinguish some vertices as more central than others and some links as more significant in that their deletion would result in larger subgraphs becoming disconnected. The obsolescence of more specific words that entered in more recent times would only affect the peripheral semantic structure.

5. Implications for brain networks

Bassette & Bullmore (2006) state that brain network architecture has likely evolved to maximize the complexity or adaptivity of function it can support while minimizing costs. Several aspects of brain structure are compatible with a selection pressure to minimize wiring costs. However, it is evident that the complete minimization of wiring would allow only local connections, leading to delayed information transfer and metabolic energy depletion. To counteract this effect, the brain also minimizes energy costs by adding several long-distance connections, creating a small-world network. The brain has likely evolved to maximize efficiency and/or minimize the costs of information processing. We may assume that small-world topology of the brain would have led to a small-world semantic configuration.

Based on Hebbian cell assemblies model by associative learning, according to which two connected neurons that fire together increase the strength of their wiring, Pulvermüller (1999, 2002) considers that strongly connected cell assemblies will form when neurons in different cortical areas are frequently active at the same time. The neuron ensembles linking phonological information and information about the actions and perceptions to which a word refers are termed word webs. They include the phonological webs in perisylvian areas, i.e. the areas next to the Sylvian fissure which runs horizontally below Brodmann areas 44 and 45 and above 22, and, in addition, neurons in more widespread cortical areas critically involved in processing perceptions and actions, and, additional neurons in various cortical sites where sensory and action-related information converges and is being merged. Visual and action associations of words are mapped by functional webs extending over perisylvian language areas and additional visual- and action-related areas in the temporo-occipital and front-central areas respectively. Words with strong semantic associations with both objects and actions (multimodal semantics), i.e. polysemous words are realized as strong connections of particularly widespread and large cortical neuronal assemblies. This argument can be extended to words referring to stimuli perceived through other modalities – ordor, taste, pain, touch, sound and color.

Ogura (1996) shows that metaphoric transfer of a lexeme from one sensory modality to another, one of the most common types of metaphoric transfer in languages, begins in frequent words first among synonymous words, and forms a polysemous word. Synonyms may be realized cortically by functional webs largely overlapping in their semantic, mainly extra perisylvian, part. The best activated word web, whose internal connection strength is likely influenced by the word frequency, would ignite first (Pulvermüller 2002). Nerve cells of the cerebral cortex are arranged in clusters, each cluster corresponding to a column of the cerebral cortex. An excited cluster projects on to other columns in the cerebral cortex by association fibers so that there is sequential activation of cluster to cluster (Eccles 1977). We may assume that the best activated word web may converge with a cluster of cells from another modality, forming a small-world neural network.

Brain function depends on adaptive self-organization of large-scale neural assemblies, but little is known about quantitative network parameters governing these processes in humans. Our quantitative analysis of structural connection patterns in semantic networks provides insights into the functioning of neural architectures in the human brain.

6. Conclusion

We have presented a quantitative study of the graph structure of the set of nouns and verbs of WordNet (version 2) and *TOE* to understand the evolution of the global organization of the English lexicon. We have demonstrated that polysemous links have a profound impact on the organization of the semantic graph, producing a highly clustered and compact graph, a small world network. The words with higher frequency and therefore with higher number of meanings construct the higher level of the hypernymy tree within each lexical category. This architecture is robust through the times, forming the basis of the small-world network. We have also suggested that the brain network architectures have evolved to maximize the complexity while minimizing costs and enhanced the evolution of the configuration of semantic structure.

Sources

TOE = *A Thesaurus of Old English.* Ed. by Jane Roberts, Christian Kay & Lynne Grundy. 1995. Centre for Late Antique and Medieval Studies, King's College London. Internet version 2008 ⟨http://libra.englang.arts.gla.ac.uk/oethesaurus⟩

WordNet, Version 2. 2003. Cognitive Science Laboratory, University of Princeton.

References

Barabási, Albert-Laszlo & Reka Albert. 1999. "Emergence of Scaling in Random Networks". *Science* 286.509–512.

Bassett, Danielle S. & ED Bullmore. 2006. "Small-World Brain Networks". *The Neuroscientist* 12.512–523.

Eccles, John C. 1977. *The Understanding of the Brain*. 2nd. New York: McGraw-Hill.

Fellbaum, Christiane. ed. 1998. *WordNet: An Electronic Lexical Database*. Cambridge, Mass. & London: The MIT Press.

Healey, Antonette DiPaolo. 2011. "Old English '*hēafod*' 'head': A Lofty Place?" *Poetica* [Tokyo] 75.29–48.

Hurford, James R. 2003. "The Language Mosaic and its Evolution". *Language Evolution* ed. by Morten H. Christiansen & Simon Kirby, 38–57. Oxford: Oxford University Press.

Hurford, James R. 2007. *The Origins of Meaning.* Oxford: Oxford University Press.

Ogura, Mieko. 1996. "Lexical Diffusion in Semantic Change: With Special Reference to Universal Changes". *Folia Linguistica Historica* 16.29–73.

Ogura, Mieko & William S-Y. Wang. 2008a. "Evolution of the Global Organization of the Lexicon". *The Evolution of Language* ed. by Andrew D. M. Smith, Kenny Smith & Ramon Ferrer i Cancho, 243–250, Singapore: World Scientific.

Ogura, Mieko & William S-Y. Wang. 2008b. "Dynamic Dialectology and Social Networks". *Selected Papers from the Fourtheenth International Conference on English Historical*

Linguistics (ICEHL 14), Begamo, Geo-historical Variation, 21–25 August 2006 ed. by Marina Dossena, Richard Dury & Mmaurizio Gotti, *vol. III,* 131–151, Amsterdam & Philadelphia: John Benjamins.

Pulvermüller Friedemann. 1999. "Words in the Brain's Language". *Behavioral and Brain Sciences* 22.253–336.

Pulvermüller Friedemann. 2002. *The Neuroscience of Language.* Cambridge: Cambridge University Press.

Sigman, Mariano & Guillermo A. Cecchi. 2002. "Global Organization of the Wordnet Lexicon". *Proceedings of the National Academy of Sciences* 99.1742–1747.

Tomassello, Michael, Malinda Carpenter, Josep Call Tanya Behne & Henrike Moll. 2005. "Understanding and Sharing Intentions: The Origins of Cultural Cognition". *Behavioral and Brain Sciences* 28.675–735.

Watts, Duncan J. 1999. *Small Worlds*. Princeton: Princeton University Press.

Watts, Duncan, J. & Steven H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks". *Nature* 393.440–442.