A LETTERS JOURNAL EXPLORING
THE FRONTIERS OF PHYSICS

Società Italiana
di Fisica

edp sciences

IOP Institute of Physics

LETTER

The role of syntax in the formation of scale-free language networks

To cite this article: Mu Yang and Haitao Liu 2022 *EPL* **139** 61002

View the article online for updates and enhancements.

You may also like

- AC and DC electrical properties of graphene nanoplatelets reinforced epoxy syntactic foam
  Ephraim Zegeye, Scott Wicker and Eyassu Woldesenbet

- Finite element simulation and experimental verification of quasi-static compression properties for 3D spacer fabric/hollow microspheres reinforced three phase composites
  Lingjie Yu, Xiaoyi He, Fanchao Liang et al.

- Mechanical properties of AlSi10MnMg matrix syntactic foams filled with lightweight expanded clay particles
  A Szlancsik, D Kincses and I N Orbulov

This content was downloaded from IP address 111.32.80.198 on 06/09/2022 at 03:21

A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

September 2022

EPL, **139** (2022) 61002
doi: 10.1209/0295-5075/ac8bf2

www.epljournal.org

The role of syntax in the formation of scale-free language networks

MU YANG[1(a)] and HAITAO LIU[2,3,4(b)]

[1] *College of Chinese Language and Culture, Nankai University - CHN-300350 Tianjin, China*
[2] *Institute of Quantitative Linguistics, Beijing Language and Culture University - CHN-100083 Beijing, China*
[3] *CLAL, Guangdong University of Foreign Studies - CHN-510420 Guangzhou, China*
[4] *Department of Linguistics, Zhejiang University - CHN-310058 Hangzhou, China*

received 5 April 2022; accepted in final form 23 August 2022
published online 5 September 2022

Abstract – The overall structure of a network is determined by its micro features, which are different in both syntactic and non-syntactic networks. However, the fact that most language networks are small-world and scale-free raises the question: does syntax play a role in forming the scale-free feature? To answer this question, we build syntactic networks and co-occurrence networks to compare the generation mechanisms of nodes, and to investigate whether syntactic and non-syntactic factors have distinct roles. The results show that frequency is the foundation of the scale-free feature, while syntax is beneficial to enhance this feature. This research introduces a microscopic approach, which may shed light on the scale-free feature of language networks.

Copyright © 2022 EPLA

Introduction. – For decades, linguists and language network researchers have accepted the view that language systems are networks with complex structures [1–4]. Network analysis has been applied to solve such numerous practical problems as language origin [3], language classification [5], language acquisition [6], disambiguation [7], measuring text complexity [8], and text classification [9]. Empirical studies have shown that nearly all language networks are small-world and scale-free [3–6,10–17]. However, this commonality fails to take into account the differences between networks. In language acquisition, for example, the scale-free, small-world features of language networks emerge along with syntax [6], and syntax is closely linked to the formation of the scale-free feature. Nevertheless, it is noteworthy that not all scale-free language networks are syntactic, such as co-occurrence networks.

The question of whether syntax has any effect on the emergence of local and global properties of language networks needs to be solved. To explore the role of syntax, ref. [16] constructs a syntactic network and two corresponding random networks, and ref. [17] builds three syntactic networks of coordinating structures with different annotation schemes. It turns out that there is no significant difference in global properties between syntactic and non-syntactic networks, and the global properties of the syntactic network are due to the frequency characteristics of language rather than syntax [16]. Therefore, the role of syntax cannot be interpreted from the global perspective only.

But it is undeniable that syntax does play a role in forming the scale-free feature. For example, ref. [18] argues that the degree of a word in a language network is equivalent to its frequency. However, ref. [4] finds that though both Zipf curves [19,20] and degree distributions obey a power law, their exponents are different (the global properties of the two types of networks in ref. [16] also differ slightly). The former is non-syntactic, while the latter is derived from syntactic networks. Thus, ref. [4] believes that this distinction is caused by syntax.

In a syntactic network, a node can be linked with a word, and the degree of a node corresponds to its valency, which can be explained as the combining ability of a word. Diverse words lead to diverse nodes, which provides a new perspective for investigating the formation of the scale-free feature. Reference [21] compares the characteristics of three central nodes in syntactic networks and explains their difference in frequency and syntactic importance. After examining the local and global properties of verb nodes, ref. [22] concludes that valency may be the primary mechanism that affects the global properties of nodes.

Micro features determine the overall structure of networks [21], and the scale-free, as well as small-world features, are generated by the interaction among nodes.

(a) E-mail: hyym@mail.nankai.edu.cn
(b) E-mail: htliu@163.com (corresponding author)

Syntactic networks and non-syntactic networks have different degrees even if their nodes are the same. So why are both syntactic networks and non-syntax networks scale-free? How are nodes generated? These questions have rarely been addressed in previous research.

In the current study, we first trace the process of converting words into nodes, then investigate the relationship between frequency distributions and degree distributions, and finally compare syntactic networks with co-occurrence networks. Following these steps, this paper aims to clarify the role of syntax in the formation of the scale-free feature. To be specific, networks are constructed based on four treebanks.

The next section presents the treebanks adopted, the methods used to build networks, and the relevant network indicators. The third section reproduces the process of converting frequency distributions into degree distributions. The fourth section analyses the role of syntax in the formation of degree distributions. The last section is the conclusion.

Resources and methods. – Dependency treebanks are a corpus with dependency syntactic annotation [23] and they are often adopted as a resource for building language networks. We used four treebanks from the Surface-Syntactic Universal Dependencies (SUD) [24] database. The English and Chinese parts of the Parallel Universal Dependencies (PUD) are retrieved for they are built on news and wikis, a genre whose style is similar to the written language. The English one includes 1000 sentences and 21176 tokens, and the Chinese one includes 1000 sentences and 21415 tokens. The spoken English treebank we used is the Reddit part (GUMReddit) of the Georgetown University Multilayer Corpus, which has 895 sentences and 15993 tokens. It is based on a blog community, Reddit, whose genre is similar to the spoken language. And we use the Mandarin part of the HK treebank which was built on film subtitles and legislative proceedings of Hong Kong, as our spoken Chinese corpora. This treebank contains 1004 sentences and 9874 tokens.

Different languages [25] and genres [26] have different syntactic expressions, and the four treebanks can minimise the influence of languages and genres as much as possible. The format of treebanks is CoNLL-X [27], and the required content extracted from it can be seen in table 1.

When building a network, it is common to convert words into nodes [3–6,11–17,21,22]. To reduce the impact of heterosemies and homographs, tuples consisting of a lemma and a word class are converted to nodes, and relations between words to edges. A syntactic network and a co-occurrence network based on table 1 can be seen in fig. 1.

For consistency, the frequencies of tuples whose lemmas are not punctuations are counted. Furthermore, tuples deliver the information about words, so they are abbreviated as words for convenience.

Table 1: Annotation of a sentence extracted from the treebanks. "ID" is the liner position of the word in the sentence. "UPOS" is the word classes. "HEAD" is the liner position of the governor of the word. "DEPREL" is the dependency relation.

ID	FORM	LEMMA	UPOS	HEAD	DEPREL
1	maybe	maybe	ADV	5	mod
2	the	the	DET	4	det
3	dress	dress	NOUN	4	compound
4	code	code	NOUN	5	subj
5	was	be	AUX	0	root
6	too	too	ADV	5	mod
7	stuffy	stuffy	ADJ	5	comp:pred
8	.	.	PUNCT	5	punct

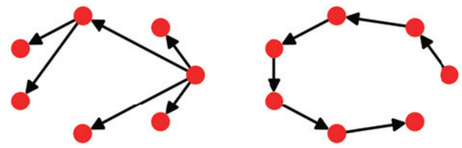

Fig. 1: Examples of a syntactic network and a co-occurrence network. On the left is the syntactic network, and on the right is the co-occurrence network.

Three Python libraries are used in this paper: *NetworkX* [28] for network construction, *scikit-learn* [29] for data fitting, and *SciPy* [30] for statistical testing.

The most basic indicator of a node is its degree, *i.e.*, the total number of edges that a node has. In a directed network, the degree can be divided into the in-degree and the out-degree. The in-degree of a node is the number of directed edges that it receives, and the out-degree of a node is the number of directed edges that it sends out. In a dependency syntactic network, these indicators possess linguistic significance.

The Probabilistic Valency Pattern Theory (PVPT) [25] suggests that the valency of words contains centrifugal and centripetal forces, that is, the ability that words have to govern other words as well as the ability to be governed by other words. It is also proposed by the theory that the valency is probabilistic. Figure 2 reveals the Probabilistic Valency Pattern (PVP) of a word or word class.

In fig. 2, G_1, G_2, $G_3, \ldots,$ G_{n-2}, G_{n-1}, G_n are words or word classes that govern the word or word class in the rectangle; namely, they are governors of the word or word class. D_1, D_2, $D_3, \ldots,$ D_{n-2}, D_{n-1}, D_n are words or word classes that are governed by the word or word class in the rectangle; namely, they are dependents of the word or word class. The arrows represent the dependency relations accompanied by their probabilities.

The PVP can be realized in a syntactic network. With the word or word class in the rectangle being regarded as a node, G's and D's are the other nodes that have the

Table 2: The power exponents γ and determination coefficients R^2 of distributions of four syntactic networks.

	English				Chinese			
	Frequency	In-degree	Out-degree	Degree	Frequency	In-degree	Out-degree	Degree
γ	2.091	2.095	2.195	2.149	2.268	2.284	2.194	2.229
	2.081	2.087	2.083	2.072	2.276	2.280	2.191	2.212
R^2	0.984	0.982	0.984	0.989	0.996	0.991	0.995	0.995
	0.987	0.987	0.990	0.987	0.989	0.987	0.983	0.981

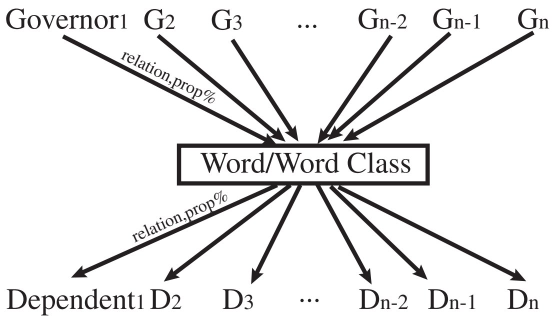

Fig. 2: The PVP of a word or word class.

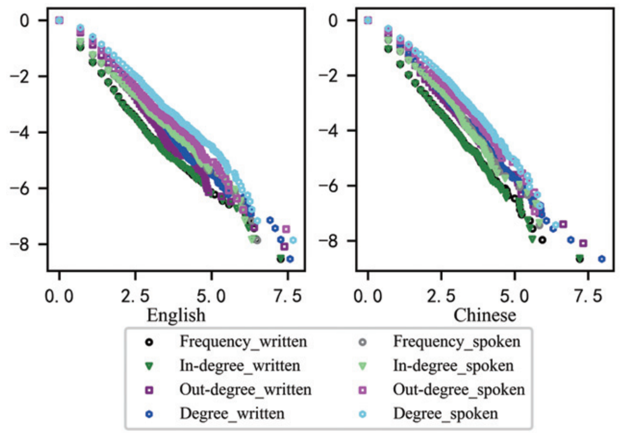

Fig. 3: Cumulative distributions based on four treebanks.

directed edges to the node, the arrows are the directed edges, and the probabilities are the weights of the edges. Hence, the degree represents the valency of a word or word class, the out-degree indicates the ability of a word or word class to govern other words, and the in-degree signifies the ability of a word or word class to be governed by other words.

The degree distribution $P(k)$ is the probability that a randomly selected node has degree k. In a directed network, the degree distribution can also be classified into the in-degree and out-degree distributions, with the former being the probability that a randomly selected node has in-degree k and the latter the probability that a randomly selected node has out-degree k.

The degree distribution of a real-world network usually obeys a power law, *i.e.*, $P(k) \sim k^{-\gamma}$. The degree distribution of a language network usually has a noise fat tail. To get better fitness, we adopt the cumulative degree distribution [31], which refers to the probability that a randomly selected node has a degree greater than or equal to k:

$$P(k) = \sum_{k'=k}^{\infty} p(k'). \tag{1}$$

If the distribution obeys a power law, then its cumulative distribution function approximately conforms to a power law with a power exponent of $\gamma - 1$:

$$P(k) = C\sum_{k'=k}^{\infty} k'^{-\gamma} \approx C\int_{k}^{\infty} k'^{-\gamma}\mathrm{d}k' = \frac{C}{\gamma-1}k^{-(\gamma-1)}. \tag{2}$$

According to eq. (2), $P(k) \sim k^{-(\gamma-1)}$.

Converting the frequency distribution into the degree distribution. – In a syntactic network, the degree of a node comes first from the frequency of its representing words, and second from the valency of these words. The degree is positively correlated with both two factors.

Initially, we plot the cumulative distributions of frequency, in-degree, and out-degree in fig. 3.

The parameters of the distributions in fig. 3 can be found in table 2. For each parameter, there are two rows of data, of which the upper one is from the written language network and the lower one is from the spoken language network (this rule also applies in table 3). We can see that neither in-degree distributions nor out-degree distributions are the same as the frequency distribution. What we want to do is to eliminate their distinction as much as possible, and the difference between frequency distributions and in-degree distributions will be erased first.

Analysis of nodes in a syntactic network based on dependency relations should consider the dependency grammar theory. The PVPT suggests that a word in a sentence can have any number of dependents, and no governor (when the word is the root of the sentence) or one governor. It might be demonstrated with a node in a syntactic network: when one occurrence of a word is extracted as a node, the out-degree of the node can be greater than or equal to 0, and the in-degree can be 0 (when the occurrence of the word is the root of the sentence) or 1. Figure 4 presents the four cases that are distinguished by the above rules.

Table 3: The power exponents γ and determination coefficients R^2 of distributions of four co-occurrence networks.

	English				Chinese			
	Frequency	In-degree	Out-degree	Degree	Frequency	In-degree	Out-degree	Degree
γ	2.091	2.099	2.073	2.089	2.268	2.274	2.258	2.236
	2.081	2.087	2.064	2.040	2.276	2.301	2.266	2.207
R^2	0.984	0.981	0.983	0.986	0.995	0.994	0.994	0.995
	0.990	0.992	0.989	0.991	0.989	0.988	0.991	0.981

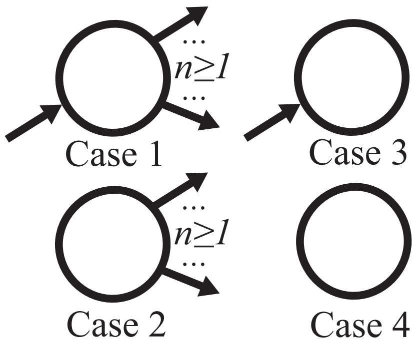

Fig. 4: Four scenarios of one occurrence of a word concerning a syntactic network formation.

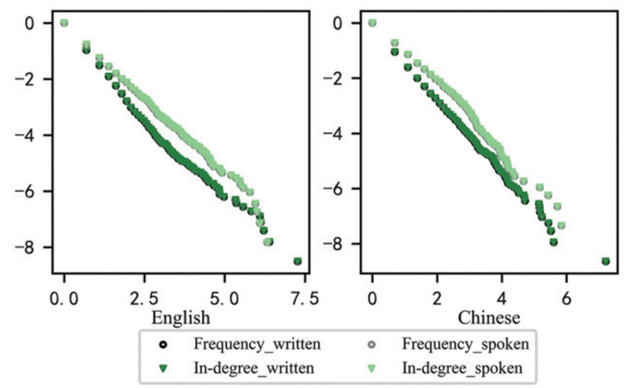

Fig. 5: Adjusted cumulative frequency distributions and cumulative in-degree distributions based on four treebanks.

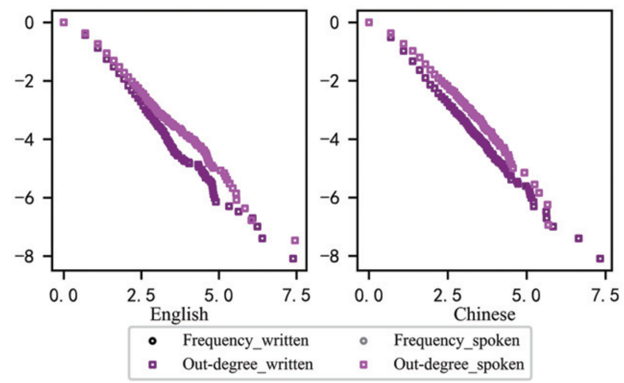

Fig. 6: Adjusted cumulative frequency distributions and cumulative out-degree distributions based on four treebanks.

In case 1, the node has in-degree 1 and out-degree n, *viz.*, this occurrence of the word has one governor and one or more dependents. In case 2, the node has in-degree 0 and out-degree n, that is, this occurrence of the word has no governor and one or more dependents. In case 3, the node has in-degree 1 and out-degree 0, that is to say, this occurrence of the word has one governor and no dependent. In case 4, the node has in-degree 0 and out-degree 0, in other words, this occurrence of the word constitutes a single-word sentence.

When all the occurrences of a word are accumulated, a node is generated. The occurrences with cases 1 and 3 present the total frequency and the total in-degree of the node. However, the occurrences for cases 2 and 4 display the total frequency of the node, but not the total in-degree. Accordingly, occurrences of cases 2 and 4 reveal the distinction between the frequency distribution and the in-degree distribution.

We then exclude the occurrences of cases 2 and 4 and compare the frequency distributions and in-degree distributions again. The results in fig. 5 state that the adjusted frequency distributions are parallel to the in-degree distributions in table 2.

Our reason for excluding the occurrences with case 2 is that some nodes with the out-degree may have no in-degree, considering that syntax allows some words to govern others but not to be governed by others. The reason for excluding occurrences of case 4 is that there may be some isolated nodes in networks since some words can syntactically form single-word sentences. Hence, it is reasonable to conclude that in-degree distributions are frequency distributions that are impacted by syntax.

Measuring out-degree distributions in this way requires some correction. For example, occurrences of cases 3 and 4 are excluded because they do not help with the total out-degree of nodes. Correspondingly, further correction is necessary because the frequency distributions are not equal to the out-degree distributions.

In cases 1 and 2, $n \geq 1$, which implies that once n is greater than 1, the total out-degree of a node will surpass its total frequency. These extra parts of out-degrees also lead to the distinction between frequency distributions and out-degree distributions. As shown in fig. 6

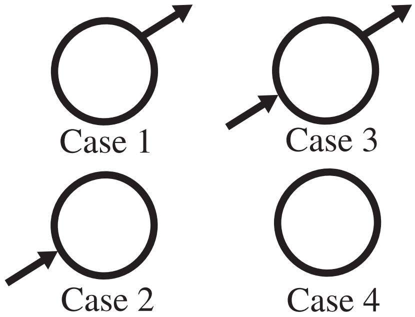

Fig. 7: Four scenarios of one occurrence of a word concerning a co-occurrence network formation.

the frequency distributions adjusted by adding these extra parts are congruent with the out-degree distributions in table 2.

The difference between the in-degree distributions and the out-degree distributions is possibly attributed to the fact that the in-degree range of one frequency of a word is less than its out-degree range. The former is $\{0, 1\}$, and the latter is $\{0, 1, 2, \ldots, +\infty-2, +\infty-1, +\infty\}$. The distinction in the range is raised by the valency, specifically, the ability that a word in a sentence can have at most one governor, but any number of dependents.

Likewise, the degree distributions can also be obtained in this way. We summarize the above two kinds of adjusted frequencies of words to obtain their degrees and get the adjusted frequency distributions that equal the degree distributions in table 2.

Consequently, it is reasonable to hold the view that the degree distribution in a weighted directed syntactic network is the frequency distribution that is impacted by syntax. In the formation of the scale-free feature of a syntactic network, frequency plays a fundamental role, and syntax plays an influential part.

There might be two possibilities for the role of syntax: one is that syntax is conducive to enhancing the scale-free feature, and the other is that syntax is adverse to enhancing the feature. The question of which of these cases holds will be answered in the next section.

The nodes of a co-occurrence network can also be analysed by the above method. A node in a co-occurrence network comes from the frequency and co-occurrence relations. According to the definition proposed in the second section, four cases of one occurrence of a word are displayed in fig. 7.

The node in case 1 has in-degree 0 and out-degree 1, for that this occurrence of the word is the first word of a sentence. The node in case 2 has in-degree 1 and out-degree 0, that is, this occurrence of the word is the last word of a sentence. In case 3, node has in-degree 1 and out-degree 1, which means that this occurrence of the word is in the middle of a sentence. In case 4, node has in-degree 0 and out-degree 0; *i.e.*, this occurrence of the word constitutes a single-word sentence.

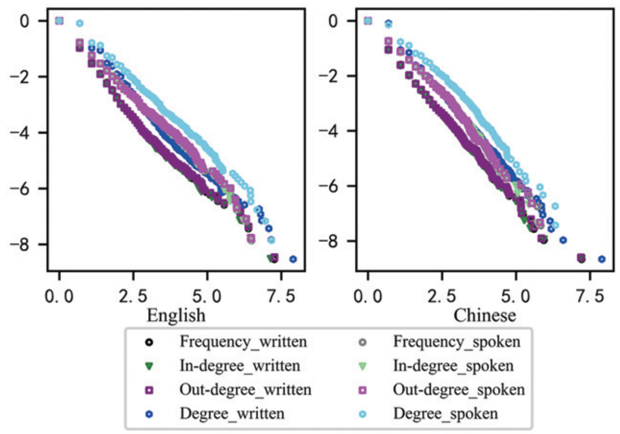

Fig. 8: Cumulative distributions based on four treebanks.

The four types of cumulative distributions for co-occurrence networks are plotted in fig. 8. Based on the frequency distributions, the in-degree distributions can be obtained by excluding occurrences of cases 1 and 4, and the out-degree distributions can be achieved by eliminating occurrences with cases 2 and 3. Consequently, the degree distributions of all four cases can be calculated.

In a co-occurrence network, co-occurrence relations play a role similar to that of syntax in a syntactic network. The degree distribution in a weighted directed co-occurrence network is the frequency distribution impacted by co-occurrence relations. Two possibilities also exist in the role of co-occurrence relations.

The frequency plays a fundamental role in both syntactic networks and co-occurrence networks. Syntax plays a part in syntactic networks, while so do co-occurrence relations in co-occurrence networks.

As fig. 3 and fig. 6 show, there is a remarkable resemblance between frequency distributions and in-degree distributions in syntactic networks. In contrast, frequency distributions bear a resemblance to both in-degree and out-degree distributions in co-occurrence networks. It indicates that despite the similar roles played by syntax and co-occurrence relations, there are still some differences between them.

The role of syntax in the scale-free feature formation. – The third section argues that the degree in syntactic networks is a combination of frequency and syntax. Hence, it is inaccurate to use the degree of a node to describe the valency of a word, since a word with a low valency and a high frequency will also have a high degree.

Mathematically, the valency of a word W can be formulated as follows:

$$valency(W) = \frac{1}{n}\sum_{i=1}^{n}|V_i|. \tag{3}$$

In this function, n is the frequency of W in treebanks. V_i is the number of the i-th W's governors and depen-

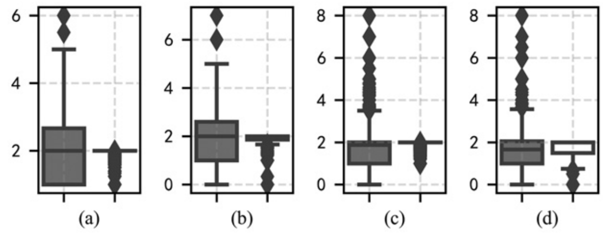

Fig. 9: Boxplots of the valency and the co-occurrency in treebanks. (a) Written English. (b) Spoken English. (c) Written Chinese. (d) Spoken Chinese.

dents. This function excludes the frequency and obtains a purely syntactic indicator, *valency*.

The ability of a word to co-occur with others can be defined as *co-occurrency*. Similarly, the co-occurrency of a word W can be formulated as:

$$\textit{co-occurrency}(W) = \frac{1}{n}\sum_{i=1}^{n}|A_i|. \qquad (4)$$

In this function, n is the frequency of W in treebanks. A_i is the number of the i-th W's adjacent words.

Notably, neither valency distributions nor co-occurrency distributions obey the power law with the determination coefficients R^2 which are all less than 0.1. We plotted boxplots of valency and co-occurrency, as shown in fig. 9.

The Wilcoxon test in fig. 9 shows that the mean of the valency ($M_a = 1.974$, $SD_a = 0.967$; $M_b = 1.970$, $SD_b = 0.973$; $M_c = 1.821$, $SD_c = 0.963$; $M_d = 1.799$, $SD_d = 0.977$) is significantly or marginally significantly ($p_{a,b,c} < 0.01; p_d < 0.75$) larger than that of the co-occurrency ($M_a = 1.868$, $SD_a = 0.300$; $M_b = 1.835$, $SD_b = 0.338$; $M_c = 1.865$, $SD_c = 0.308$; $M_d = 1.820$, $SD_d = 0.429$). Moreover, the standard deviation implies that the co-occurrency is distributed more uniformly than the valency. The data show that the co-occurrencies have a weaker influence on degree distributions than the valencies.

The above findings are anticipated. As a matter of fact, the four cases in fig. 7 are contained in fig. 4. The range in which the valency can vary is much greater than that in which the co-occurrency can vary. In addition, case 3 in fig. 7 is dominant, which leads to the high concentration of the co-occurrency around 2. Therefore, the impact of the co-occurrency on degree is insignificant, while the valency promotes the emergence of the differences among nodes.

Then, we discuss the specific impact of the valency on the degree distribution and the formation of the scale-free feature.

In a scale-free network, there are a small number of nodes with remarkably large degrees and a large number of nodes with fairly low degrees. The degree distribution of a scale-free network obeys a power law, and the range of its degree is greater than $(\langle k\rangle - \langle k\rangle^{\frac{1}{2}}, \langle k\rangle + \langle k\rangle^{\frac{1}{2}})$ [32]. Compared with the number of nodes with a degree in this range, there are fewer nodes with a degree greater than $\langle k\rangle + \langle k\rangle^{\frac{1}{2}}$, and more nodes with a degree less than $\langle k\rangle - \langle k\rangle^{\frac{1}{2}}$.

Table 4: The average degree of degree distributions based on frequency distributions and the number of nodes in three segments.

	English		Chinese	
	Written	Spoken	Written	Spoken
$\langle k\rangle$	3.692	5.562	3.191	4.755
$(\langle k\rangle + \langle k\rangle^{\frac{1}{2}}, \langle k\rangle + \infty)$	428	266	596	233
$(\langle k\rangle - \langle k\rangle^{\frac{1}{2}}, \langle k\rangle + \langle k\rangle^{\frac{1}{2}})$	1507	298	1453	319
$(\langle k\rangle - \infty, \langle k\rangle - \langle k\rangle^{\frac{1}{2}})$	3136	2037	3752	1159

Table 5: The rank-frequency distribution table with valencies of the PUD treebank.

Rank	Word	Frequency	Valency
1	(the, DET)	1441	1.000
2	(be, AUX)	651	3.007
3	(of, ADP)	599	2.003
4	(in, ADP)	501	2.023
5	(and,CCONJ)	456	1.000
...	...	...	...
5070	(Caesar, PROPN)	1	1.000
5071	(proconsul, NOUN)	1	2.000

Table 6: Proportion of strong valencies in three segments.

	English		Chinese	
	Written	Spoken	Written	Spoken
$(\langle k\rangle + \langle k\rangle^{\frac{1}{2}}, \langle k\rangle + \infty)$	84.1%	81.2%	72.3%	76.0%
$(\langle k\rangle - \langle k\rangle^{\frac{1}{2}}, \langle k\rangle + \langle k\rangle^{\frac{1}{2}})$	76.5%	78.1%	70.2%	67.4%
$(\langle k\rangle - \infty, \langle k\rangle - \langle k\rangle^{\frac{1}{2}})$	58.1%	64.5%	51.5%	54.8%

Frequency distributions satisfy the above conditions, which means that the frequency distributions are scale-free. The number of nodes in the three segments can be seen in table 4.

Finally, we calculate the valencies of the words represented by nodes in these three segments one by one. Without changing the position of words in tables of the rank-frequency distributions, we add their valencies, as shown in table 5.

The valencies of very few words are less than 1 (these words constitute single-word sentences), and the valencies of the other words are greater than or equal to 1. Valencies equal to 1 are regarded as weak valencies, while valencies greater than 1 as strong ones. Statistics on the proportion of strong valencies in the three segments are listed in table 6.

Table 6 illustrates that the proportion in $(\langle k\rangle + \langle k\rangle^{\frac{1}{2}}, \langle k\rangle + \infty)$ is the largest, and that in $(\langle k\rangle - \infty, \langle k\rangle - \langle k\rangle^{\frac{1}{2}})$ is the smallest. It suggests that a word with a high frequency tends to have a strong valency. This feature is helpful to diversify the differences among nodes and enhance the scale-free feature of networks.

Conclusion. – It is widely acknowledged that almost all real-world networks are small-world and scale-free. Like things in the real world, the nodes of a real-world network have diverse characteristics that are crucial to the generation of the overall structure of the network. Through investigating the generation mechanism of nodes, this paper argues that both the syntactic network and the co-occurrence network are scale-free, due to the fundamental role of the frequency. This finding may be useful in real applications such as authorship attribution [33] and language classification [34,35], where taking both syntactic and non-syntactic factors into account can lead to better results.

Given the prevalence of the scale-free, small-world features, there is no significant difference in the global organization of different networks [16,17]. The process of forming the scale-free feature can be illustrated in detail from the perspective of nodes. The data validates that syntax is more conducive to widening the differences among degrees and forming the scale-free feature than co-occurrence relations.

Furthermore, this paper demonstrates that it is feasible to analyse the network from a microscopic perspective, which facilitates the explanation of language networks from the linguistic point of view. Due to space limitations, this paper only discusses scale-freeness, and the equally important small-worldness will be studied from the perspective of the motif or others in our subsequent research.

REFERENCES

[1] Lamb S. M., *Pathways of the Brain: The Neurocognitive Basis of Language* (John Benjamins Publishing Company, Amsterdam) 1999.
[2] Hudson R., *Language Networks: The New Word Grammar* (Oxford University Press, Oxford) 2007.
[3] Solé R. V., *Nature*, **434** (2005) 289.
[4] Liu H., *Physica A*, **387** (2008) 3048.
[5] Abramov O. and Mehler A., *J. Quant. Linguist.*, **18** (2011) 291.
[6] Jiang J., Yu W. and Liu H., *Front. Psychol.*, **10** (2019) 291.
[7] Silva T. C. and Amancio D. R., *EPL*, **98** (2012) 58001.
[8] Amancio D. R. *et al.*, *EPL*, **100** (2012) 58002.
[9] de Arruda H. F., Costa L. D. F. and Amancio D. R., *EPL*, **113** (2016) 28007.
[10] Watts D. J. and Strogatz S. H., *Nature*, **393** (1998) 440.
[11] Barabási A.-L. and Albert R., *Science*, **286** (1999) 509.
[12] Ferrer i Cancho F. and Solé R. V., *Proc. R. Soc. London, Ser. B*, **268** (2001) 2261.
[13] Ferrer i Cancho F., Solé R. V. and Köhler R., *Phys. Rev. E*, **69** (2004) 051915.
[14] Amancio D. R., Silva F. N. and Costa L. D. F., *EPL*, **110** (2015) 68001.
[15] Chen H., Chen X. and Liu H., *PLoS ONE*, **13** (2018) e0192545.
[16] Liu H. and HU F., *EPL*, **83** (2008) 18002.
[17] Liu H., Zhao Y. and Huang W., *Glottometrics*, **20** (2010) 38.
[18] Masucci A. P. and Rodgers G. J., *Phys. Rev. E*, **74** (2006) 026102.
[19] Zipf G. K., *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology* (Addison-Wesley Press, Cambridge, Mass.) 1949.
[20] Bian C. *et al.*, *EPL*, **113** (2016) 18002.
[21] Chen X. and Liu H., in *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*, edited by Mehler A. *et al.* (Springer, Berlin, Heidelberg) 2016, pp. 187–201
[22] Čech R., Mačutek J. and Žabokrtský Z., *Physica A*, **390** (2011) 3614.
[23] Mel'čuk I. A., *Dependency Syntax: Theory and Practice* (State University of New York Press, Albany) 1988.
[24] Gerdes K. *et al.*, *Proceedings of the Second Workshop on Universal Dependencies* (Association for Computational Linguistics, Brussels) 2018, pp. 66–74.
[25] Yan J. and Liu H., *Chinese Lexical Semantic Workshop* (Springer, Cham) 2021, pp. 152–162.
[26] Poiret R. and Liu H., *J. Fr. Lang. Stud.*, **30** (2020) 355.
[27] Buchholz S. and Marsi E., *Proceedings of the Tenth Conference on Computational Natural Language Learning* (Association for Computational Linguistics, New York City) 2006, pp. 149–164.
[28] Hagberg A. A., Schult D. A. and Swart P. J., *Proceedings of the 7th Python in Science Conference, Pasadena, CA, USA, 2008*, pp. 11–15.
[29] Pedregosa F. *et al.*, *J. Mach. Learn. Res.*, **12** (2011) 2825.
[30] Virtanen P. *et al.*, *Nat. Methods*, **17** (2020) 261.
[31] Newman M. E. J., *SIAM Rev.*, **45** (2003) 167.
[32] Barabási A.-L., *Network Science* (Cambridge University Press, Cambridge) 2016.
[33] Akimushkin C. *et al.*, *Physica A*, **495** (2018) 49.
[34] Liu H. and Li W., *Chin. Sci. Bull.*, **55** (2010) 3458.
[35] Liu H. and Xu C., *EPL*, **93** (2011) 28005.