Classifying Syntactic Categories in the Chinese Dependency Network

Xinying Chen
Xi'an Jiaotong University
School of International Study
China
xy@yuyanxue.net

Haitao Liu
Zhejiang University
Department of Linguistics
China
lhtzju@gmail.com

Kim Gerdes
Sorbonne Nouvelle
ILPGA, LPP (CNRS)
France
kim@gerdes.fr

Abstract

This article presents a new approach of using dependency treebanks in theoretical syntactic research: The view of dependency treebanks as combined networks. This allows the usage of advanced tools for network analysis that quite easily provide novel insight into the syntactic structure of language. As an example of this approach, we will show how the network approach can provide an interesting angle to discuss the degree of connectivity of Chinese syntactic categories, which it is not so easy to detect from the original treebank.

1 Hierarchical Features inside Language

It is a widely accepted idea that language is a complex, multi-level system (Kretzschmar 2009, Beckner et al. 2009, Hudson 2006, Mel'čuk 1988, Sgall 1986, Lamb 1966). Languages can be described and analyzed on different linguistics levels, such as morphology, syntax, and semantic etc. Moreover, these different linguistics levels form a surface-deep hierarchy (Mel'čuk 1981). Besides the macro multi-level hierarchy of languages, the unequal relationships between linguistic units in sentences are also widely recognized by linguists. Such as the concept of governor in dependency grammar, head of phrase in HPSG etc. In this article, we aim to define a new kind of one-directional asymmetrical relationships between linguistic units, half-way between the macro-model of language and the syntactic analysis of single sentences.

Hierarchies have been recognized as one of the key features of any formal language description on two very different levels:

Firstly, linguistics as a whole wants to describe the relation between Saussure's signified and signifier (Saussure 2011) (or Mel'čuk's meaning and text (Mel'čuk 1981), or Chomsky's logical and phonetic structure (Chomsky 2002)). Although the theories differ widely on how the steps between the two sides of language should be described, all theories developed a hierarchy of interrelated structures that build up the language model.

Secondly, each subdomain of linguistics has developed hierarchical structures describing each utterance, for example on a semantic, communicative, phonological, and, most noteworthy, syntactic level.

It is important to reflect on the wide gap between these two types of hierarchies: One describing the language as a whole (i.e. all languages), the other just describing one utterance of one particular language by hierarchical means. This paper describes how intermediate structures can be discovered, intermediate in a sense that they describe a global feature of syntax of one language, which could then be compared to equivalent analyses of other languages.

In sections 2 to 4, we will show that syntactic categories of a language as a whole are related in complex ways, thus establishing a hierarchy among the categories. In order to proceed to the actual analysis we first have to show two points:

1. The notion of syntactic category (or part of speech, POS) has an existence in the syntactic model as a whole that goes beyond the classification of individual words.
2. A dependency treebank provides means of studying meaningful relationships between syntactic categories.

To 1: When developing a system of categorization for a given language, the syntactician already has a global view of grouping together syntactic units that have comparable distributional or morphological properties with the goal to allow for the expression of rules that generalize beyond the actual linguistic evidence. However, the analysis remains local in

Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015), pages 74–81,
Uppsala, Sweden, August 24–26 2015.

a sense that the syntactician does not create relationships inside the proposed categorization, the objective of the analysis simply being to put forward distinctive features that can be tested and applied to the data. It is thus reasonable to search for ways of exploring general properties that have been implicitly encoded with the categorization.

To 2: The aforementioned distributional and morphological properties of syntactic categories make them an ideal candidate in the search for global syntactic feature of language, but the theoretical aspects and the generalizability at the basis of the categorization are difficult to study empirically. Syntactic dependency, however, describes links that represent the distributional properties of a word: Words of the same category are in general part of a paradigm of words that can hold the same syntactic position. A dependency treebank can accordingly be seen as relations between paradigms of words.

2 Networks

Over the last decade or so, driven by theoretical considerations as well as by the simple availability of large amount of connected data, network analysis has become an important factor in various domains of research ranging from sociology, biology to physics and computer science (Barabási & Bonabeau 2003, Watts & Strogatz 1998).

Equally, digital language data and the popularity of statistical approaches had the first effect that many linguists, who are mainly interested in theoretical questions as well as NLP researchers have started to quantitatively describe microscopic linguistic features in a certain level of a language system by using authentic language data. Despite the fruitful findings, one question remains unclear. That is, how can the statistical analysis of raw texts (e.g. n-gram based language models) or of treebanks (syntactic models, i.e. the statistical prediction of likely syntactic relations) provide linguistic insight? Or put differently, how does a complete empirical language system look like?

As an attempt to answer this question, the network approach, an analysis method emphasizing the macro features of linguistic structures, has been introduced into linguistic studies (Solé 2005, Ferrer-i-Cancho & Solé 2001). By analyzing different linguistics networks constructed from authentic language data, many linguistic features, such as lexical, syntactic or semantic features have been discovered and successfully applied in linguistic typological studies thus revealing the huge potential of linguistic networks research (Cong & Liu 2014).

What is particularly interesting about the recent development in this area is that researchers have been able to systematically analyze linguistic features beyond the sentence level since the network approach is not intrinsically limited by traditional linguistic feature annotations in corpora based on the lexical or the sentence level. It seems possible that linguistic network model, as the representation of the whole body of language data, is a better approach to explore the human language systems.

Moreover, just as all the networks constructed based on real data (Barabási & Bona-beau 2003, Watts & Strogatz 1998), the linguistic networks are ‘small world’ and ‘scale free’ networks too (Solé 2005, Ferrer-i-Cancho & Solé 2001, Liu 2008), which indicates that there are central nodes (Chen & Liu 2015, Chen 2013), or hubs, in language networks. And that will provide a natural hierarchy between the nodes or the units of the networks.

3 Building a Syntactic Network

When we talk about the structure of languages, the first thing that naturally comes to our mind is the syntactic structure. Both phrase structure grammar and dependency grammar have been developed and deployed in the analysis of corpora. In the past decade, dependency annotated treebanks have become the latest hype in empirical linguistics studies. Driven by the statistical NLP development and the linguist’s fascination of creating a treebank following specific theoretical principles, considerable efforts have been devoted to treebank creation and analysis (among many others Marcus et al. 1993, Lacheret et al. 2014, Mille et al. 2013). Solid theoretical foundation and available well-annotated data made syntactic structural analysis the candidate of choice for most studies in linguistic network analysis just as in the present study.

In more detail, dependency treebanks, especially multi-layer dependency treebanks such as Ancora-UPF, offer interesting connections between texts and the representation of meaning, which allow us to pursue further discussion about the semantic structure more easily in the future. In addition, since our goal is finding the hierarchy between linguistic units of the same type, phrase structure, which introduces different levels of constituents, is less apt for the task than dependency structure.

Dependency treebanks commonly encode two kinds of information for each word: the word's syntactic relation with its governor and the word's syntactic category (or POS). Thus, a dependency treebank can be seen as a collection of dependency trees on words or on POS tags. We will call the first a 'word dependency tree' and the latter a 'POS dependency tree' which will be the base of the present experiment. Both trees can represent the syntactic structure of linguistic units in a sentence, while POS trees are more abstract and less detailed in a way.

Various previous research has been undertaken on the network analysis of syntactic dependency treebanks (Chen & Liu 2011, Chen et al. 2011, Čech et al. 2011, Liu 2008, Ferrer-i-Cancho 2005), some also based on the same Chinese dependency treebank used for this study (Liu 2008, Chen 2013, Chen & Liu 2011). These approaches all used word dependency trees, thus obtaining results on the network behavior of individual words. The central nodes in networks based on word dependency trees, however, are highly correlated with the frequency of the word itself and it is difficult to account for the influence of the unequal distribution of the different words. In POS dependency trees, the different classes are more evenly distributed and the role of frequency of categories may be less crucial.

Moreover, the high number of different word types makes the data exploration and explanation more complex than in networks based on POS dependency trees. Our specific goal of this present study is to find the hierarchies on Chinese categories (or POS) in the syntactic network which is constructed on empirical language data, or more specifically, the Chinese dependency treebank.

The basic idea underlying dependency networks is very simple: Instead of viewing the trees as linearly aligned on the sentences of the corpus, we fuse together each occurrence of the same POS to a unique node, thus creating a unique and connected network of POS, in which the POS are the vertices and dependency relations are the edges or arcs. This connected network is then ready to undergo common network analysis with tools like UCINET (Borgatti et al. 2002), PAJEK (Nooy et al. 2005), NETDRAW (Borgatti 2002), CYTOSCAPE (Shannon 2003), and so on. For more details, we refer to Liu (2008) for a description of multiple ways of network creation from dependency treebanks.

For the present work, we used the following treebank of Chinese, the XBSS treebank (Liu 2008): The XBSS has 37,024 tokens and is composed of 2 sections of different styles:

- "新闻联播" xin-wen-lian-bo 'news feeds' (name of a famous Chinese TV news program), is a transcription of the program. The text is usually read and the style of the language is quite formal. The section contains 17,061 words.
- "实话实说" shi-hua-shi-shuo 'straight talk' (name of a famous Chinese talk show), is of more colloquial language type, containing spontaneous speech appearing in interviews of people of various social backgrounds, ranging from farmers to successful businessmen, The section contains 19, 963 words.

Both sections have been annotated manually as described by Liu (2006). Table 1 shows the file format of this Chinese dependency treebank,

Sentence Order	Dependent			Governor			Dependency type
	Order	Character	POS	Order	Character	POS	
S1	1	*zhe*	pronoun	2	*shi*	verb	subject
S1	2	*shi*	verb	6	。	punctuation	main governor
S1	3	*yi*	numeral	4	*ge*	classifier	complement of classifier
S1	4	*ge*	classifier	5	*zuqiu*	noun	attributer
S1	5	*zuqiu*	noun	2	*shi*	verb	object
S1	6	。	punctuation				

Table 1. Annotation of a sample sentence.
这是一个足球 *zhe-shi-yi-ge-zu-qiu* 'this is a football'

which is similar to the CoNLL dependency format, although a bit more redundant (double information on the governor's POS) to allow for easy exploitation of the data in a spreadsheet and converting to language networks. The data can be represented as simple dependency graphs as shown in Figure 1: 1a is the dependency tree of the words in the sentence and 1b illustrates the dependency relationship between POS in this example. The trees both show a bottom-top hierarchy between the linguistic units in this sample sentence.

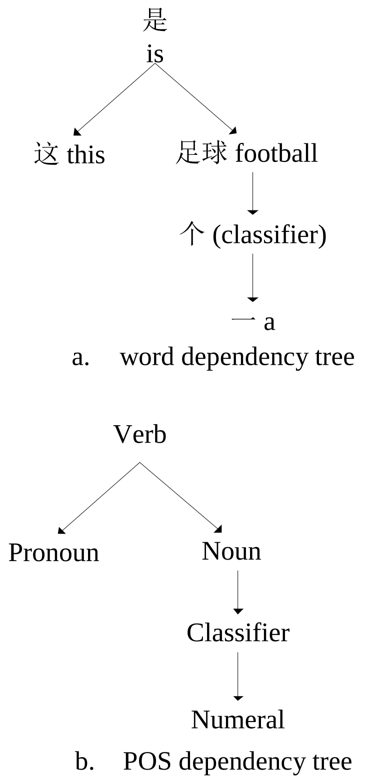

Figure 1. The graph of the dependency analysis of 这是一个苹果 *zhe-shi-yi-ge- zu-qiu* 'this is a football'

With POS as nodes, dependencies as arcs, and the frequency of the dependencies as the value of arcs, we can build a network. For example, our Chinese treebank can be represented as Figure 2, an image, generated by the network analysis software Pajek, which gives a broad overview of the global structure of the treebank (excluding punctuation).

The resulting network it is a fully connected network without any isolated vertices. As we set the distance between POS inversely proportional to the value of arcs (the detailed information of arcs values can be found in the table of appendix C), the graph actually can give us an intuitive idea of the 'clusters' of syntactic connections between POS already.

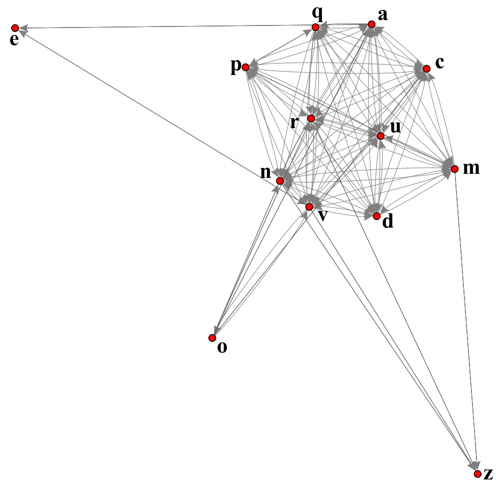

Figure 2. The POS network of the treebank. The details of all codes and symbols in tables and figures in this paper are available in Appendix A.

For minimizing the effect of genre difference to the data result, we chose to include two similar size sections of text in our treebank. However, some other factors may remain that could possibly affect the result of the study, such as the size of the treebank, the annotation schema, the language type, etc. We will leave these discussions for further work.

The reason we chose Chinese rather than other 'big' languages such as English, French or Spanish is that Chinese, as an isolating language, lacks morphological changes. Since there is no 'difference' between tokens and lemmas in Chinese dependency treebanks, Chinese syntactic networks built on dependency treebanks would only have one unique form for each treebank while every single inflectional language would have two different types of syntactic networks, word-type syntactic network and lemma syntactic network. As so, Chinese is a better choice for this study considering no ambiguity of defining a 'syntactic network'.

4 Data Analysis

There are two simple ways in a network model to detect the hierarchy of nodes. First by the degrees which represents the number of different types of links one node can have; second by the summed value of arcs which indicates, we believe, the intensity of the combination capacity of one node has. When one node can link to more nodes (or has a higher degree), as well as more connections to other nodes (or summed value of arcs), it is more likely to be the ‘hub’ or occupying a central position of the network structure. When we analyze or visualize a network, software such as Pajek try to optimize the positions of nodes so that they will fit the distance difference between pairs of nodes. However, for more precise result, we need to do a multi-dimensional scaling (MDS) analysis. With Ucinet (V 6.186), we did a non-metric MDS analysis to our POS network data, and made the network data a two dimensional perceptual map as in Figure 3. The actual coordinate values of all the nodes are listed in Table 2.

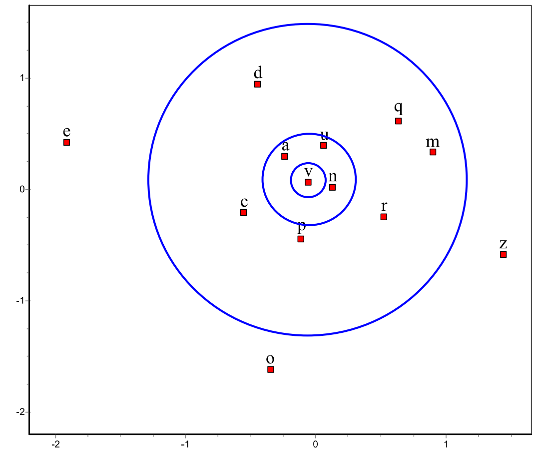

Figure 3. The perceptual map of the network.

Kruskal (1964) proposed to measure the quality of MDS result by index STRESS (the equation of STRESS can be found in appendix B). When the STRESS index is no more than 0.1, the result is acceptable for further discussion. The STRESS index of our analysis here is 0.100, which means that we are good to continue.

	POS	y	x
n	**noun**	0.021	0.127
v	**verb**	0.066	-0.059
r	**pronoun**	-0.244	0.520
q	**classifier**	0.615	0.633
m	**numeral**	0.334	0.897
p	**preposition**	-0.448	-0.115
a	**adjective**	0.297	-0.238
z	**affix**	-0.581	1.439
u	**auxiliary**	0.395	0.059
d	**adverb**	0.946	-0.447
c	**conjunction**	-0.204	-0.555
o	**mimetic word**	-1.619	-0.347
e	**interjection**	0.422	-1.913

Table 2. The coordination of POS in figure 3.

According to Figure 3, we can roughly divide the POS in to central, middle, and marginal parts. Since we are talking about the syntactic dependency structure here, verbs are expected be the very center of syntactic structures. With verb as the center, nouns, adjectives, and auxiliaries constructed scattered closely around the verb and constructed as the central part of the diagram, mimetic words, interjections, and affixes are far away from the center and they are the marginal part of the diagram. All the others POS fell between these two extremes and become the middle part of the diagram. The hierarchical structure of POS seems relatively clear according to the perceptual map already.

Yet, for more accurate result, we rely on the coordinate values of the POS in Figure 3 to do a clustering analysis, see Figure 4 (done with OriginPro, V 9.0). The result further confirmed the division we did according to Figure 3 but in greater details. Such as, we can find ‘smaller groups’ inside the central and middle parts of the network:

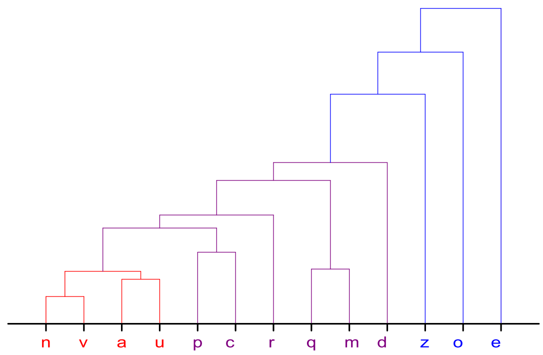

Figure 4. The clustering analysis result.

- Inside the central part, there are actually two small groups: verbs and nouns, adjectives and auxiliaries.
- Inside the middle part, there are also two closely tied small groups: propositions and coordinators, numerals and classifiers.

All these results correspond surprisingly well to our understanding of the Chinese language. For example, verbs are for sure the very center of the syntactic structure just as illustrated in Figure 3. Nouns, auxiliaries and adjectives are relatively frequent words in the treebank and hold important roles in syntactic well-formed sentences, they form the central part and are thus located in a relatively higher position in the POS hierarchy we built and showed in Figure 3 and Figure 4. Meanwhile, the infrequent mimetic words, interjections, and affixes are syntactically not very important in Chinese, therefore they have been put on a lower position, a more marginal part, of our POS hierarchy. Theoretically, the POS hierarchy may be caused by the uneven distribution of valence of POS, or more generally, by the unequal capacity of combination force of the POS. The bigger the valence a POS has, i.e. the stronger its capacity of combination it owns, the higher possibility of getting into the central part of the syntactic system.

When we look into the resulting data, it seems that the word or POS frequency played a role here. It seems that the more frequent POS in the treebank has been put in the more central part in the hierarchy, see table 3.

POS		Frequency
n	**noun**	11, 014
v	**verb**	9, 562
r	**pronoun**	3, 411
u	**auxiliary**	3, 195
d	**adverb**	2, 634
a	**adjective**	1, 976
q	**classifier**	1, 491
p	**preposition**	1, 244
m	**numeral**	1, 561
c	**conjunction**	903
z	**affix**	413
e	**interjection**	3
o	**mimetic word**	1

Table 3. The frequency distribution of POS .

As much as connections between our results and the POS frequency, they are not fully corresponding to each other, such as:

- nouns have the highest frequency in *XBSS* but they are not in the most central position in the hierarchy while verbs are.
- pronouns have the third highest frequency but only belong to the middle part of the system, meanwhile the adjectives locate on the relatively central position with a moderate frequency.
- conjunctions have relatively low frequency but they locate on a position closer to the center than numerals, classifiers, and adverbs do, and these POS all have greater frequency than conjunctions do.

We think the frequency of POS might be an explicit result of constructing sentences by following the rules of the Chinese syntactic system, which is a fully connected system that has a hierarchical feature, see Figure 2. The frequency distribution index treats the linguistic units as individuals while the network model also address the importance of the connections between linguistic units.

Although further discussion is needed for understanding the connections between the frequency distribution of POS and the positions that POS occupies in syntactic network, we speculate that the hierarchy feature may be a motive behind the POS frequency distribution or word frequency distribution, rather than, contrarily, that the central position is due to the high frequency.

5 Conclusion

For a long time, the discussion of the hierarchical features of language is mainly focusing on the hierarchical structure between different linguistic layers or inside a sentence. It seems that there is an empty gap between the very detailed sentence structures and general linguistic layers. If we find hierarchical structure inside a sentence as well as the text-meaning process, then cannot we find hierarchical structures in between, inside each linguistic layer?

The challenge of breaking the boundary of sentences while remaining reasonable syntactic structures was met by the network model. With the dependency treebank, we constructed a POS network and did several quantitative analysis to the language network data.

With empirical data support, our study found a clear hierarchical structure of POS in Chinese syntactic system. Although further study is needed for a more insightful discussion, our preliminary results made us believe that the hierarchical configuration is a natural (i.e. inborn or core) feature of language systems, which can be seen not only in the hierarchy of different linguistics levels but also inside certain linguistics layer. Moreover, such configurations probably exist inside each linguist level.

The study showed a method that not only allows us to do quantitative analysis on language data, but also empowers the theoretical discussion by offering support of concrete empirical data. We can discuss the hierarchy features of language by analyzing the authentic language data and visually present it to give us a more intuitive understanding of abstract concepts.

We believe the hierarchy we observed in this study can be seen as the result of the uneven distribution of linguistic units' valence, or more generally, linguistic units' capacity of combination. Since the valence of linguistic units is, actually a concept which closely links to semantics and syntax, we expect the hierarchical structure that we found in this study to equally be observable on the semantic level although classes in propositional semantics differ from syntactic categories. The common points and differences of hierarchical structures between syntactic and semantic layers can be a possible future direction of the methods presented in this study, as soon as comparable semantic treebanks will be available.

As we mentioned before, in future work, furthermore, we have to explore the effect of some factors such as the size of the treebank, the annotation scheme, the language type, etc.

This paper addresses the importance of developing techniques of treebank exploitation for syntactic research ranging from theorem verification to discovery of new linguistic relations invisible to the eye. We advocate in particular for the usage of network tools in this process and showed how a treebank can, and, in our view, should be seen as a unique network.

Acknowledgments

This work was supported in part by the National Social Science Fund of China (11& ZD188).

References

Barabási A L. and Bonabeau E. 2003. Scale-free networks. *Scientific American*, 288(5), 50-9.

Beckner C, Blythe R, Bybee J, Christiansen MH, Croft W, Ellis NC, Holland J, K JY, Larsen-Freeman D, Schoenemann T. 2009. Language is a complex adaptive system: Position paper. *Language learning*, 59(s1), 1-26.

Borgatti S P. 2002. *NetDraw: Graph visualization software*. Analytic Technologies, Harvard.

Borgatti S P, Everett M G, Freeman L C. 2002. *Ucinet for Windows: Software for social network analysis*. Analytic Technologies, Harvard.

Čech R, Mačutek J, Žabokrtský Z. 2011. The role of syntax in complex networks: Local and global importance of verbs in a syntactic dependency network. *Physica A*, 390(20), 3614-3623.

Chen X. 2013. Dependency Network Syntax. In *Proceedings of DepLing 2013*, 41-50.

Chen X, Liu H. 2015. Function nodes in the Chinese syntactic networks. In *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks. Series on Understanding Complex Systems*, Springer.

Chen X, Liu H. 2011. Central nodes of the Chinese syntactic networks. *Chinese Science Bulletin*, 56(1): 735-740.

Chen X, Xu C, Li W. 2011. Extracting Valency Patterns of Word Classes from Syntactic Complex Networks. In *Proceedings of DepLing 2011*, 165-172.

Chomsky N. 2002. *Syntactic structures*. Walter de Gruyter.

Cong J, Liu H. 2014. Approaching human language with complex networks. *Physics of life reviews*, 11(4), 598-618.

De Saussure F. 2011. *Course in general linguistics*. Columbia University Press.

Deschenes L A, David A. 2000. Origin 6.0: Scientific Data Analysis and Graphing Software. *Journal of the American Chemical Society*, 122(39), 9567-9568.

Ferrer i Cancho R. 2005. The structure of syntactic dependency networks: insights from recent advances in network theory. Problems of quantitative linguistics, 60-75.

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

Hudson R. 2006. *Language Networks: The New Word Grammar*. Oxford University Press.

Kretzschmar W A. 2009. *The linguistics of speech*. Cambridge University Press.

Kruskal J B. 1964. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29(2), 115-129.

Lacheret A, Kahane S, Beliao J, Dister A, Gerdes K, Goldman J P, Obin N, Pietrandrea P, Tchobanov A. 2014. Rhapsodie: a Prosodic-Syntactic Treebank for Spoken French. In *Language Resources and Evaluation Conference*.

Lamb S. 1966. *Oueine Of Stratificational Grammar*. Washington: Georgetown University Press.

Liu H. 2008. The complexity of Chinese dependency syntactic networks. *Physica A*, 387, 3048-3058.

Liu H. 2006. Syntactic Parsing Based on Dependency Relations. *Grkg/Humankybernetik*, 47:124-135.

Mel'čuk I. 1988. *Dependency Syntax: Theory and Practice*. Albany: State University of New York Press.

Mel'čuk I. 1981. Meaning-Text Models: Arecenttrendin Sovietlinguistics. *Annual Review of Anthropology*, 10, 27-62.

Mille S, Burga A, Wanner L. 2013. AnCoraUPF: A Multi-Level Annotation of Spanish. In *Proceedings of DepLing 2013*, 217-226.

Nooy W, Mrvar A, Batagelj V. 2005. *Exploratory Network Analysis with Pajek*. Cambridge University Press, New York.

Sgall P, Hajičová E, Panevová J. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Dordrecht: Reidel Publishing Company.

Shannon P, Markiel A, Ozier O, Baliga N S, Wang J T, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11), 2498-2504.

Solé R. 2005. Syntax for free? *Nature*, 434, 289.

Watts D. J. and Strogatz S. H. 1998. Collective dynamics of 'small-world' networks. *nature*, 393(6684), 440-442.

Appendix A. Codes meaning

code	meaning
a	adjective
c	conjunction
d	adverb
e	interjection
m	numeral
n	noun
o	mimetic word
p	preposition
q	classifier
r	pronoun
u	auxiliary
v	verb
z	affix

Appendix B. The equation of index STRESS

$$STRESS = \sqrt{\frac{\sum_{ij} (\delta_{ij} - d_{ij})^2}{\sum_{ij} d_{ij}^2}}$$

Appendix C. The value of arcs in the POS network

gov \ dep	n	v	r	q	m	p	a	z	u	d	c	o	e
n	3, 246	822	489	966	239	23	642	12	1, 417	30	115	0	0
v	5, 429	5, 707	1, 809	399	124	1, 098	705	1	1, 505	2049	632	1	1
r	71	12	67	15	1	2	3	361	11	6	7	0	0
q	31	15	471	16	1, 000	0	15	0	15	4	2	0	0
m	18	17	27	4	144	0	12	39	19	13	1	0	0
p	829	162	154	16	5	4	15	0	10	23	34	0	0
a	245	145	97	30	22	31	101	0	115	442	35	0	2
z	0	0	0	0	0	0	0	0	0	0	0	0	0
u	548	681	264	32	17	73	374	0	18	33	50	0	0
d	9	16	3	3	3	3	2	0	4	22	1	0	0
c	543	311	20	6	5	9	35	0	68	11	11	0	0
o	3	21	3	0	0	0	1	0	0	0	1	0	0
e	0	0	0	0	0	0	0	0	0	0	0	0	0