Using a Chinese treebank
to measure dependency distance

HAITAO LIU, RICHARD HUDSON, and ZHIWEI FENG

Abstract

This article describes a method for calculating the 'dependency distance' between the words in a text – i.e. the number of words that separate each word from the word on which it depends syntactically – and reports the results of applying this method to a Chinese treebank. This study shows that Chinese dependencies tend strongly to be governor-final and that the mean dependency distance of words is much higher for Chinese than for other languages that have been studied including English, German and Japanese. It is unclear whether this difference means that Chinese is syntactically more difficult to process.

Keywords: Dependency syntax; Chinese treebank; dependency distance.

1. Introduction

A treebank is not only useful for practical projects in computational linguistics such as training and evaluating parsers, but can also be used as a resource for quantitatively analyzing the syntactic structures of texts and drawing conclusions about how humans process the language concerned. In this paper we report on the dependency patterns in a news Chinese treebank, and discuss two sets of statistical results involving the direction and length of the dependencies. The latter measure is 'dependency distance', and allows interesting comparisons between and within languages which we shall discuss. We do not discuss either the principles of building a treebank (Abeillé 2003), nor the general relations between treebank and linguistic theories.[1]

We shall take it for granted in this paper that the syntactic structure of a sentence consists of nothing but the dependencies between individual words – an assumption that is widely accepted not only in computational linguistics but also in theoretical linguistics and that can be justified independently

(Hudson 2007). The following are generally accepted as the core properties of a syntactic dependency relation (Tesnière 1959; Hudson 1990):

1. It is a binary relation between two linguistic units.
2. It is usually asymmetrical, with one of the two units acting as the governor and the other as dependent.
3. It is labeled and the type of a dependency relation is usually indicated using a label on top of the arc linking the two units.

A further assumption that we shall make is that the linguistic units related in this way are single words rather than phrases. As for dependency distance, this is the linear distance between governor and dependent measured in terms of the number of words from one to the other. Distance is an important property of a dependency because of its implications for the cognitive costs of processing the dependency, so the average dependency distance of a text is also an important comparative measure for the light that it throws on the cognitive demands of the language concerned. In this paper, we introduce a new approach to using a treebank to measure the dependency distance and dependency directions of a language.

Section 2 briefly introduces the dependency syntax of Chinese and the dependency treebank used in this study. The usefulness of dependency distance is discussed in section 3, and some concepts and our methods for measuring dependency distance in a treebank are also given in this section. Section 4 presents the results, which we discuss in Section 5.

2.   Chinese dependency syntax and treebank

Our tentative treebank is built on the news (*xinwen lianbo*) of China Central Television, a genre which is intended to be spoken but whose style is similar to the written language. We selected four complete news broadcasts as the annotated material. The final treebank includes 709 sentences and 17363 word tokens, so the mean sentence length is 24 words.

The syntactic analysis was carried out manually. In order to make the annotation manageable, we recruited two groups of people to annotate the texts. The first group consisted of 20 undergraduates of Chinese linguistics who annotated 500 tokens each, and the other group contains a single graduate of linguistics who did the remaining 10000 tokens. Then two teachers of Chinese linguistics reviewed the entire corpus.

The analysis specifies two kinds of information about each word: its word class, and its dependency relations to other words defined in terms of a tagset of dependency types. As far as the word classes are concerned, we used a set of word classes based on the national standard "Part of speech (POS) tagset

for Chinese information processing" (2003) and the widely used "Chinese Grammar for middle-school teaching". Our system includes 13 main classes: noun (n), verb (v), adjective (a), adverb (d), pronoun (r), preposition (p), numeral (m), classifier (q), conjunction(c), interjection (e), particle (u) and onomatopoeia (o) and punctuation (bd). This set excludes some POS tags in the national standard set which do not work on the level of syntax, and differs from traditional school grammar by giving some functional (particle) words an important position in the syntax, because they often play a crucial role by determining the dependency relation between two words. Figure 1 shows the hierarchy of POS tagset in Chinese.[2]
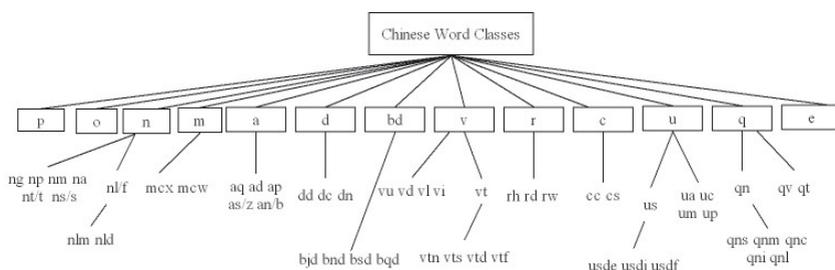

Figure 1. *A hierarchy of word classes in Chinese*

The dependency tagset contains 20 complements and 14 adjuncts. Chinese has a few more types of complement than other languages (Maxwell and Schubert 1989), because Chinese has to use functional words for signalling the grammatical functions which often are morphologically realized in other languages. The dependency types are listed in Table 1 and summarised in Figure 2.

Example (1) is a simple sentence which we can use to illustrate the analysis.

(1)   这 是 三 个 例子 。
      'This is three classifier example.' (literal translation)
      'These are three examples.' (regular translation)

Table 2 shows the analysis of (1) in terms of this dependency syntax, with each word token distinguished by a number which shows the linear order of words.

This format expresses the properties of each dependency relation, but sometimes it is also helpful to construct a connected directed labeled graph (Mel'cuk 1988: 23) such as Figure 3.

This analysis provides a suitable basis for exploring our two questions about the direction and length of dependency relations. However, our treebank is still rather small so we also converted the Chinese dependency treebank of

Table 1.   *Chinese dependency types*

| Type	Label	Type	Label
Main governor	S	Sentential object	SentObj
Subject	SUBJ	Auxiliary verb	ObjA
Object	OBJ	Coordinating mark	C-
Indirect Object	OBJ2	Adverbial	AVDA
Subobject	SUBOBJ	Verb adjunct	VA
Subject Complement	SOC	Attributer	ATR
Prepositional Object	POBJ	Topic	TOP
Postpositional Complement	FC	Coordinating adjunct	COOR
Complement	COMP	Epithet	EPA
Complement of usde '的'	DEC	Numeral adjunct	MA
Complement of usdi '的'	DIC	Aspect adjunct	TA
Complement of usdf '得'	DFC	Adjunct of sentence end	ESA
Object of Pba '把'	BaOBJ	Parenthesis	InA
Plural complement	PLC	Clause adjunct	CR
Ordinal complement	OC	Correlative adjunct	CsR
Complement of classifier	QC	Particle adjunct	AuxR
Construction of Pbei '被'	BeiS	Punctuation	Punct
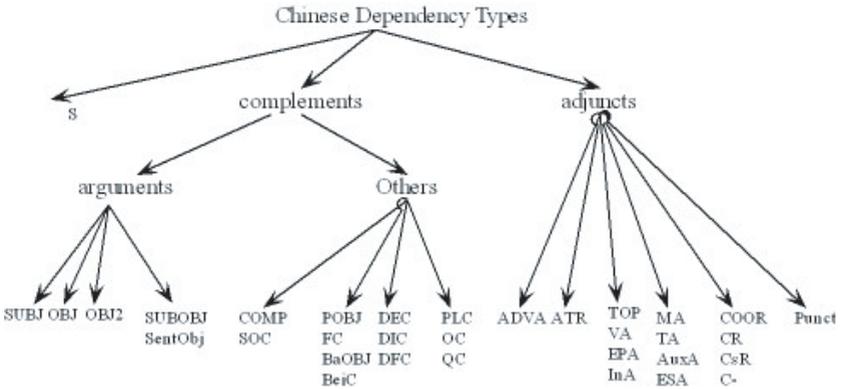
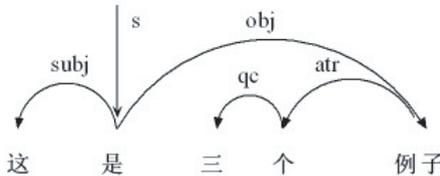Figure 2.   *A hierarchy of dependency types in Chinese*

Figure 3.   *The analysis displayed as a graph*

Table 2. *Analysis of a sample sentence*

| Dependent | | | Governor | | | Dependency |
Order	Character	POS	Order	Character	POS	Type
1	这	r	2	是	v	subj
2	是	v	–	–	–	s
3	三	m	4	个	q	qc
4	个	q	5	例子	n	atr
5	例子	n	2	是	v	obj

the Harbin Institute of Technology (HIT) into our format. The HIT treebank contains 9,996 sentences with 178,467 word tokens, which are mostly taken from newspapers – a similar genre selection to our treebank, though the mean sentence length is only 17 words, somewhat shorter than the 24 of our corpus.

3. Measuring dependency distance and dependency direction

Dependency distance is the linear distance between governor and dependent. The concept was first used in (Heringer et al. 1980: 187), who extracted the idea from the depth hypothesis of Yngve on phrase structure grammar (Yngve 1960, 1996). The term 'dependency distance' was introduced in Hudson (1995: 16) and defined as "the distance between words and their parents, measured in terms of intervening words."

Measuring dependency distance (DD) is useful for:

1. Predicting syntactic difficulty. If human parsing (i.e. syntactic analysis) of a sentence is seen as an incremental procedure for transforming a linear string into a tree or graph, a word can be removed from working memory only once it can be linked with some other word to form a dependency relation. According to Miller (1956), this memory has a limited capacity with 7 ± 2 units, which impacts on the human capacity for parsing. This question has received wide attention in current cognitive science (Gibson 1998; Gibson and Pearlmutter 1998; Grodner and Gibson 2005). Gibson (1998) hypothesizes "(1) the longer a predicted category must be kept in memory before the prediction is satisfied, the greater is the cost for maintaining that prediction; and (2) the greater the distance between an incoming word and the most local head or dependent to which it attaches, the greater the integration cost." In our case, the greater the dependency distance, the harder the processing.

2. Explaining the mechanisms of children language learning. Ninio (1998) shows that children learning both English and Hebrew combine three

122 words much more easily when dependency distance is the minimum than
123 when one word is separated from the word on which it depends.
124 3. Designing better parsing algorithms for natural language processing. Col-
125 lins proves that "the relative order of the words and the distance between
126 them will also strongly influence the likelihood of one word modifying the
127 other." (1996: 187) Buch-Kromann (2006: 100) also mentions the impor-
128 tance of dependency distance for parsing of a natural language.

129 Several of these arguments are based on the assumption that DD can be aver-
130 aged across the words in a text, and that the resulting average DD provides a
131 relevant basis for comparing different texts in a single language, or even for
132 comparing texts in different languages. If two texts in different languages are
133 otherwise comparable – e.g. they are both examples of casual conversation
134 or of scripted news broadcasts – then we may take them as representative
135 of the syntactic patterns in their respective languages and draw conclusions
136 about the languages themselves. Hudson (1995), Hiranuma (1999) and Ep-
137 pler (2004) have used manual analyses of very short texts to calculate the
138 dependency distances in English, Japanese and German. The main innova-
139 tion of our research project is that the calculation is based on a reasonably
140 large treebank.

141   How, then, should one calculate the result automatically? Formally, let
142 $W_1 \ldots W_i \ldots W_n$ be a word string. For any dependency relation between the
143 words W_a and W_b, if W_a is governor and W_b is dependent, then the depen-
144 dency distance (DD) between them can be defined as the difference *a-b*; by
145 this measure, adjacent words have a DD of 1 (rather than 0 as when DD is
146 measured in terms of intervening words). When *a* is greater than *b*, the DD is a
147 positive number, which means that the governor is after the dependent; when
148 *a* is smaller than *b*, the DD is a negative number and the governor precedes
149 the dependent. For instance, in the above example "这 是 三 个 例子 。", the
150 DD of '是—这' is $2 - 1 = 1$; '个—三' is $4 - 3 = 1$; '是—例子' is $2 - 5 = -3$;
151 '例子—个' is $5 - 4 = 1$.
152   As we shall see below in discussing the direction of dependencies, it is
153 sometimes useful to distinguish the directions of DD with positive and neg-
154 ative numbers, and this is the reason why this project, unlike other projects,[3]
155 always counts the distance between two words as at least 1. However, in mea-
156 suring dependency distance the relevant measure is the absolute dependency
157 distance (ADD), where the 'ADD' between words W_a and W_b can be defined
158 as ADD = $|a\text{-}b|$.
159   The mean dependency distance (MDD) of an entire sentence can be defined
160 as in (2)

161 (2) $MDD(the\ sentence) = \frac{1}{n-1} \sum_{i=1}^{n-1} |DD_i|$

Here *n* is the number of words in the sentence. DDi is the dependency distance of the *i*-th syntactic link of the sentence. According to this formula, the MDD of the above example is: $(1 + 1 + 1 + 3)/4 = 1.5$.

This formula can also be used to calculate the mean dependency distance of a larger collection of sentences, such as a treebank; cf. (3).

$$(3) \qquad MDD\,(the\ sample) = \frac{1}{n-s} \sum_{i=1}^{n-s} |DD_i|$$

in this case, *n* is the total number of words in the sample, *s* is the total number of sentences in the sample. DD_i is the dependency distance of the *i*-th syntactic link of the sample.

Another application is to calculate the MDD for a specific type of dependency relation in a sample, as shown in (4).

$$(4) \qquad MDD\,(dependency\ type) = \frac{1}{n} \sum_{i=1}^{n} DD_i$$

Here *n* as the number of examples of that relation in the sample. DD_i is the dependency distance of the *i*-th syntactic link in the set of that dependency type.

Dependency Distance is not the only important and interesting variable in syntactic structure. The other one which we studied is Dependency Direction, the contrast between governor-initial and governor-final dependencies. For any given dependency, of course, this is a purely qualitative difference, but for a collection of dependencies such as those in a sentence or a text it is a quantitative measure according to the relative proportions of the two directions in the collection. As explained earlier, our method of measuring Dependency Distance allows us to distinguish the two directions according to whether the position-number of the governor is higher than that of the dependent (governor-final) or lower (governor-initial). Measuring DD of a language can more clearly determine if the language is governor-final or governor-initial. This idea comes from Tesnière (1959: 22–24, 32–33), who considers such a measure useful to studying language typology. Liu (submitted) proposes a method based on dependency directions to classify a language typologically, and the experiment with 20 languages shows that the dependency direction is useful for language typological classification.

4.   Measuring dependency distance and direction in Chinese

There are 17,363 word tokens and 166,54 dependencies in our sample (treebank), most of which are between adjacent words. In the following report our results are normalised to make them comparable to the results of other projects (Hudson 1995; Hiranuma 1999; Eppler 2004), where DD is mea-

sured in terms of the number of intervening words rather than as the difference between the words' position-numbers. For example, adjacent words have a DD of 0 in the other projects instead of 1 in our original analysis, and the MDD of example (1) is (0+0+0+2)/4 = 0.5.

Table 3 shows the overall frequencies in our tree bank of the various values for ADD (the absolute dependency distance, which ignores direction).

Table 3.   *The frequency of each distance length in the Chinese treebank*

ADD	Dependencies	Cumulative total (Proportion)
= 0	9377	9377 (56.3%)
= 1	2391	11,768 (70.7%)
= 2	1246	13,014 (78.1%)
= 3	794	13,808 (82.9%)
= 4	583	14,391 (86.1%)
= 5	401	14,792 (88.8%)
= 6	353	15,145 (90.9%)
= 7	282	15,427 (92.6)
= 8	228	15,655 (94%)
= 9	173	15,828 (95%)
$\geq$ 10	826	16,654 (100%)

However, ADD interacts strongly with direction. Most dependencies (68.5%) are positive (governor-final), but adjacent dependencies are even more likely to be governor-final as can be seen in Table 4 and Table 5, where adjacent dependencies (where DD = 0) are contrasted with all others. The differences in Table 4 are highly significant ($p < 0.001$).

Table 4.   *Dependencies in our treebank*

	Positive DD (governor final)	Negative DD (governor initial)	Total
ADD = 0	7563	1815	9378
ADD > 0	3848	3428	7276
	11,411	5243	16,654

Table 5.   *Percentage of positive and negative dependencies in our treebank*

	Positive DD (governor final)	Negative DD (governor initial)	
ADD = 0	80.6%	19.4%	56.3%
ADD > 0	52.9%	47.1%	43.7%
	68.5%	31.5%	100%

209 Using the formula mentioned in Section 3, we get the MDD of posi-
210 tive/negative and all dependencies in our treebank, which is shown in Table 6.

Table 6.   *MDD and its distribution in our treebank*

	Positive DD	Negative DD	All DD
MDD	1.25	3.3	1.89

211 The HIT treebank includes 178,467 word tokens and 168,470 dependencies.
212 The related figures are presented in Table 7, Table 8 and Table 9.

Table 7.   *Dependencies in HIT treebank*

	Positive DD (governor final)	Negative DD (governor initial)	Total
ADD = 0	68802	16766	85568
ADD > 0	43321	39581	82902
	112123	56347	168470

Table 8.   *Percentage of positive and negative dependencies in HIT treebank*

	Positive DD (governor final)	Negative DD (governor initial)	
ADD = 0	80.4%	19.6%	50.8%
ADD > 0	52.3%	47.7%	49.2%
	66.6%	33.4%	100%

Table 9.   *MDD and its distribution in HIT treebank*

	Positive DD	Negative DD	All DD
MDD	1.34	2.99	1.89

213 The differences in Table 7 are also highly significant ($p < 0.001$).
214 Our figures also show a strong interaction between dependency distance
215 and the individual dependency relations such as *subject* or *object*, which can
216 be useful for understanding MDD differences between languages. The data in
217 Table 10, calculated using Formula (4), shows the MDD of some dependency
218 relations in our Chinese treebank. These figures need more research but they
219 show very clearly that some dependencies tend to be much longer than others.

Table 10. *Dependency distances and distribution for distinct dependency types in a Chinese treebank*

Dependency type	Dependencies	MDD[4]
Complements	190	−0.01
Aspect adjunct	217	−0.08
Classifier complement	467	0.11
Attribute	4625	0.55
Subject	1395	1.66
Adverbial	2375	2.44
Object	1673	−2.90
'Ba'(°Ñ) construction as object	41	4.76
Parenthesis	95	6.8
Sentential object	172	−6.81
Clausal relation	729	−8.96

5.   Discussion

Section 4 presents the results of our analysis for two Chinese dependency treebanks. In this section, we try to interpret the results and hope to make a link between dependency distance and linguistics. During the discussion, we will also compare our results with earlier works,[5] although such comparison is difficult.

Eppler (2004) calculates the dependency distance of English and German. After normalization, the distribution of English and German is as shown in Table 11 (Eppler 2004: 156–158). Eppler explains the difference in terms of the syntactic differences between the two languages. For instance, in German the main verb is often at the end of the sentence and word order is freer than in English.

Table 11. *Comparison of MDD in samples of English and German speech*[6]

	German	English
MDD	0.87	0.49
ADD = 0	490 (65%)	463 (78%)
ADD > 0	264 (35%)	133 (22%)
Total	754	596

Hiranuma (1999: 313) measures the MDD of conversational English and Japanese as 0.386 and 0.43, respectively.[7] Japanese is a head-final language, which might be expected to have a higher MDD because more dependents are separated from the head than in a language like English where the dependents tend to occur on either side of the head. However, Hiranuma reports

that conversational Japanese in fact has a similar MDD to English and suggests, as explanation, that this is because Japanese uses a smaller number of dependents than English. Hiranuma's experiment also shows that formal style may have a greater MDD.

On the basis of two dependency treebanks,[8] Ferrer i Cancho (2004) calculates that about 50%–67% of the links in sentences are formed between adjacent words and 16%–25% are formed at distance 1. However, although his treebanks contain different languages (Romanian and Czech), his method does not lead to an MDD of the two languages. In contrast, Buch-Kromann (2006) reports, in the Danish Dependency Treebank,[9] that only 44% of all dependents are immediately preceded by their governor, though 88% are fewer than 5 words apart from their governor. Based on the treebanks of 20 languages, Liu (2008a) finds almost 50% of dependency relations are built between adjacent words.

These figures suggest that different languages favour different dependency distances. On the other hand, dependency distance appears to be positively correlated with style, with casual speech favouring much shorter distances than more formal, possibly scripted, speech and writing. This being so, it is possible that some of these differences between treebanks are due to the style of the texts concerned rather than to differences between the languages. However, even if the samples for different languages are taken from different styles, the numerical differences are striking and deserve more investigation.

As for Chinese, the data in Table 3–5 show:[10]

1. Chinese is basically a governor-final language.
2. Only 56.3% of dependents are adjacent to their governors, which is a lower proportion than in English (78%) and German (65%).
3. The MDD is much greater in governor-initial relations than in governor-final.

According to the studies by Collins (1996) for English, the figures for ADD ≤ 4 are 95.6% and for ADD ≤ 9 are 99%, which are considerably higher than for Chinese (Table 3), suggesting a much larger 'tail' of high distances in Chinese than in English.

The interaction of MDD and direction in Chinese also deserves a great deal more study. According to both Tables 6 and 9, the MDD for governor-initial dependencies is more than twice as great as that for governor-final dependencies. This is clearly a feature of Chinese dependencies, but we do not know either how to generalise to other languages, or to explain it in Chinese.

The above discussion also shows that the MDD of Chinese is much greater than in English, German and Japanese.[11] If the MDD can be used as a criterion to estimate the syntactic difficulty, can we say that Chinese is a syntactically more difficult language than other languages? Or does it show that MDD is

only valid in a single language, so we can not use it to compare the syntactic difficulties of two languages?

One way to explore these questions is to control the effect of meaning and style by looking in detail at a text in one language and its translation into the other, so we selected a sample text with about 100 words in English,[12] and compared the dependencies in it and in its published Chinese translation. The English text contains 106 words and the Chinese translation 108 words. However in spite of their similar style and meanings, the MDDs of the two texts were very different: 0.51 for English and 1.42 for Chinese. It is easy to find structural explanations for these differences; for example, whereas prepositional phrases follow the modified noun in English, in Chinese they precede it, which means that the preposition's complement inevitably separates it from the modified noun. However, these explanations run counter to the common functional view that a language's structure adapts to minimise the processing difficulties of its users. We cannot rule out the possibility that Chinese makes greater demands of its users than the other languages that we have studied.

The main aim of this paper is not to answer these important theoretical questions but to show how this kind of research can raise such questions. The method that we used in the investigations reported here has the following original features:

1. It is based on dependency treebanks, which can be built using standard methods from computational linguistics;
2. It measures positive and negative dependency distance, i.e. dependency directions;
3. It can not only be used to measure the MDD of a sentence, but also to measure that of a whole language sample or of a specific type of dependency relation.

We leave it to future work to develop these methods to improve our ability to compare dependency distances across languages and to explore and explain the reasons for differences of the kind we have revealed here.

Bionotes

Haitao Liu is professor of applied and computational linguistics at Communication University of China (CUC), Beijing. His research interests include computational linguistics, corpus linguistics and dependency grammar. E-mail: lhtcuc@gmail.com

Richard Hudson is an emeritus professor of linguistics in the Department of Phonetics and Linguistics at University College London.

Zhiwei Feng is professor at Communication University of China (CUC), Beijing.

Notes

1. Six consecutive TLT (Treebanks and Linguistic Theories) workshops reveal that there is close relationship between treebanks and linguistic theories. The first "Treebanks and Linguistic Theories" workshop was held in Sozopol, Bulgaria in 2002; the second in Vaxjo, Sweden in 2003; the third in Tuebingen, Germany in 2004; the fourth in Barcelona, Spain in 2005; the fifth in Prague, Czech in 2006 and the sixth in Bergen, Norway in 2007.
2. Liuand Huang (2006) includes a detailed explanation of the subclass tags.
3. Collins (1996) and Ferrer i Cancho (2004) also define dependency distance as the order difference between governor and dependent, but they and other colleagues (Hudson 1995; Hiranuma 1999; Eppler 2004) do not distinguish the direction of a DD and only use the absolute value of the difference.
4. Here the signs ($-$) or ($+$) to indicate if the dependent is before or after the governor. For instance, if MDD has a value -2.5, which means that the dependency type is governor-initial and there are approximately 1.5 words between the governor and the dependent.
5. We limit our discussion to dependency grammar formalism, although there are also some works from cognitive science, which often are under phrase structure formalism.
6. While Eppler and Hiranuma calculate a MDD of a sentence with n words, they use the n as the divisor. The difference is very small. In our case, the MDD using n is 1.81, using $n-1$ is 1.89.
7. English: 1035 words; Japanese: 2117 words.
8. Romanian with 21,275 words and 2340 sentences, Czech with 563,067 words and 31,701 sentences.
9. 100,200 words, 5540 sentences.
10. Liu (2008b) has similar conclusions based on 5 Chinese treebanks with different genres and annotation schemes.
11. According to Liu (2008a), Chinese has also the greatest dependency distance in 20 languages, which are Chinese, Japanese, German, Czech, Danish, Swedish, Dutch, Arabic, Turkish, Spanish, Portuguese, Bulgarian, Slovenian, Italian, English, Romanian, Basque, Catalan, Greek, Hungarian.
12. The text is extracted from Pinker's "The Language Instinct" (Perennial Classic, 2000). Dependency analysis at http://www.phon.ucl.ac.uk/home/dick/enc/pinker.htm. The Chinese translation is from the Chinese version of this book (Shantou University Press, 2004).

References

Abeillé, Anne (ed.)
    2003   *Treebank: Building and using parsed corpora*. Dordrecht: Kluwer.
Buch-Kromann, Matthias
    2006   *Discontinuous grammar. A dependency-based model of human parsing and language acquisition*. Copenhagen: Copenhagen Business School Dr. ling.merc. dissertation.
Collins, Michael
    1996   A new statistical parser based on bigram lexical dependencies. *Proceedings of the Association for Computational Linguistics* 34. 184–191.

Eppler, Eva
    2004   *The syntax of German-English code-switching*. London: University College
            London unpublished PhD.
Ferrer i Cancho, Ramon
    2004   Euclidean distance between syntactically linked words. *Physical Review E* 70.
            056135.
Gibson, Edward
    1998   Linguistic complexity: locality of syntactic dependencies. *Cognition* 68(1).
            1–76.
Gibson, Edward & Neal J. Pearlmutter
    1998   Constraints on sentence comprehension. *Trends in Cognitive Sciences* 2(7).
            262–268.
Grodner, Daniel & Edward Gibson
    2005   Consequences of the serial nature of linguistic input for sentential complexity.
            Cognitive Science 29(2). 261–290.
Heringer, Hans-Jürgen, Bruno Strecker & Rainer Wimmer
    1980   *Syntax: Fragen-Lösungen-Alternativen*. Munich: Wilhelm Fink Verlag.
Hiranuma, So
    1999   Syntactic Difficulty in English and Japanese: A textual study. *UCL Working
            Papers in Linguistics* 11. 309–322.
Hudson, Richard
    1990   *English Word Grammar*. Oxford: Blackwell.
Hudson, Richard
    1995   Measuring syntactic difficulty. Unpublished paper.
            http://www.phon.ucl.ac.uk/home/dick/difficulty.htm
Hudson, Richard
    2007   *Language networks: The new Word Grammar*. Oxford: Oxford University Press.
Liu, Haitao
    2008a   Dependency distance as a metric of language comprehension difficulty. *Journal
            of Cognitive Science* 9(2). 159–191.
Liu, Haitao
    2008b   A quantitative study of Chinese structures based on dependency treebanks.
            Yangtze River Academic 3. 120–128.
Liu, Haitao
    submitted. Dependency direction as an index of linguistic typology: A method based on
            dependency treebanks.
Liu, Haitao & Wei Huang
    2006   A Chinese dependency syntax for treebanking. In *Proceedings of the 20th
            Pacific Asia Conference on Language, Information and Computation*, 126–
            133. Wuhan, Beijing: Tsinghua University Press.
Maxwell, Dan & Klaus Schubert
    1989   *Metataxis in practice: Dependency syntax for multilingual machine translation*.
            Dordrecht: Foris.
Mel'cuk, Igor A.
    1988   *Dependency syntax: Theory and practice*. Albany: State University Press of
            New York.
Miller, George
    1956.   The magical number seven plus or minus two: Some limits on our capacity for
            processing information. *Psychological Review* 63. 81–97.

Ninio, Anat.
    1998   Acquiring a dependency grammar: The first three stages in the acquisition of
                multiword combinations in Hebrew-speaking children. In G. Makiello-Jarza,
                J. Kaiser & M. Smolczynska (eds.). *Language acquisition and developmental*
                psychology. Crakow: Universitas.
Tesnière, Lucien
    1959   *Eléments de la syntaxe structurale.* Paris: Klincksieck.
Yngve, Victor.
    1960   A model and a hypothesis for language structure. *Proceedings of the American*
                Philosophical Society 104(5). 444–466.
Yngve, Victor
    1996   *From grammar to science: New foundations for general linguistics.* Amsterdam
                & Philapdelpia: John Benjamins.