Menzerath-Altmann Law in Syntactic Dependency Structure

Ján Mačutek
Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics
Department of Applied Mathematics and Statistics
Slovakia
jmacutek@yahoo.com

Radek Čech
University of Ostrava
Faculty of Arts
Department of Czech Language
Czech Republic
cechradek@gmail.com

Jiří Milička
Charles University, Prague
Faculty of Arts
Institute of Comparative Linguistics, and
Institute of the Czech National Corpus
Czech Republic
jiri@milicka.cz

Abstract

According to the Menzerath-Altmann law, there is a relation between the size of the whole and the mean size of its parts. The validity of the law was demonstrated on relations between several language units, e.g., the longer a word, the shorter the syllables the word consists of. In this paper it is shown that the law is valid also in syntactic dependency structure in Czech. In particular, longer clauses tend to be composed of shorter phrases (the size of a phrase is measured by the number of words it consists of).

1 Introduction

Some language properties can be considered a result of general mechanisms influencing human language behaviour. The mechanisms can be expressed by language laws which can have, in the ideal case, a form of a mathematic formula. The mathematical formalization allows to test the validity of a law statistically, and, in addition, it opens a door towards building a theory, i.e., a system of interconnected valid laws (see, e.g., Bunge, 1967; Altmann, 1978, 1993). In this paper, a particular instance of language laws, namely, the Menzerath-Altmann law (MAL hereafter) in syntactic dependency structure is scrutinized. The MAL (Cramer, 2005a) is, in general, a law expressing a mechanism which controls mutual relations between sizes of language units belonging to "neighbouring" language levels (e.g., between lengths of words and syllables, clauses and words, etc.), see Section 2 for details. Our aim is to test the validity of the MAL in syntactic dependency structure; namely, we hypothesize that the relation between the size of the clause and the mean size of its parts (i.e., phrases; for details, see Section 3) follows the MAL. If the hypothesis is corroborated, syntactic dependency structure can be included among other linguistic "domains" which are substantially influenced by the very general mechanism expressed by the MAL. Consequently, in such a case the general status of the MAL in language is confirmed (and strengthened), and some fundamental properties of syntactic dependency structure can be seen (and possibly explained) from a new point of view.

The article is organized as follows. The MAL is introduced in Section 2 (with some basic examples). Section 3 describes the methodology applied in this study. The language material from which data are extracted is presented in Section 4. Section 5 summarizes the results

Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017), pages 100-107,
Pisa, Italy, September 18-20 2017

achieved. Finally, the paper is concluded by Section 6, where also perspectives for future research are pointed to.

2 Menzerath-Altmann law

The MAL speaks, in general, about the relation between sizes of a construct and its constituents. It is named after two linguists: Paul Menzerath, who observed length of German words and length of syllables which the words consist of, and Gabriel Altmann, who contributed to a substantial generalization of the law.

The verbal formulation of the law changed over time. Its first version (the longer the word, the shorter syllables in the word, see Menzerath, 1954) was a description of the relation between length of words and syllables. The current version of the MAL (Altmann, 1980) is more general, expressing a relation between sizes of two language units which are "neighbours" in the language unit hierarchy, such as syllables and words, sentences and clauses, etc. (the greater the whole the smaller its parts). We note that the hierarchy of the units is a nested structure[1] (e.g., a sentence consists of clauses, which consist of words, which consist of syllables, which consist of phonemes[2]). Thus, one usually speaks about constructs and constituents (e.g., words and syllables). Furthermore, the formulation of the MAL from Altmann (1980) is not so strict with respect to the monotonicity of the relations between lengths of a construct and its constituents. In some cases, constituent's length does not achieve its maximum in constructs with length one, but its peak is shifted to the right. Hence, the MAL can be presented in its most general form as "the mean size of constituents is a function of the size of the construct".

The mathematical formula corresponding to the abovementioned general verbal expression of the MAL is

(1) $$y(x) = ax^b e^{-cx},$$

with $y(x)$ being the mean size of constituents if the size of the construct is x; a, b, c are parameters[3]. However, in many cases its special case of (1) for $c = 0$, i.e.,

(2) $$y(x) = ax^b,$$

fits data sufficiently well[4]. This special case describes a strictly decreasing trend of the constituent size. The goodness of fit is usually evaluated in terms of the determination coefficient R^2 (the higher R^2, the better fit). It is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(S_i - y(i)\right)^2}{\sum_{i=1}^{n}(S_i - \bar{S})^2},$$

where S_i is the observed mean size of constituents for constructs of size i, $\bar{S}$ is the mean of values S_i, $i = 1, 2, \dots, n$, and $y(i)$ are theoretical values from a model (which is given by (1) or (2) in this paper). A model is usually considered good enough if it achieves $R^2 \geq 0.9$, see Mačutek and Wimmer (2013).

The validity of the MAL was corroborated on relations between pairs of several language units in many languages (language material from both dictionaries and texts was used, i.e., the MAL seems to be valid for both types and tokens). We mention several examples which cover relations among some traditional language units[5]. Kelih (2010) investigated the relation between word length in syllables and syllable length in graphemes in Serbian. Gerlach (1982) chose word (in German) as the construct as well, but he measured word length in the number of morphemes (with morpheme length determined in the number of phonemes). Teupenhayn and Altmann (1984) showed that the MAL can be used also to describe the relation between sentence length (in clauses) and clause length (in words). An example of this relation, data from a German text together with a curve corresponding to the theoretical model of the MAL, can be seen in Figure 1. The data and the curve displayed in the figure can be considered typical for the MAL.

[1] In fact, there are several parallel nested structures. If word is taken as a construct, both syllables and morphemes can serve as its constituents; depending on whether one works directly with a written text or with its phonological or phonetic transcription, the size of a syllable can be measured in the number of graphemes, phonemes or sounds, etc. The choice of language units is conditioned by the technical tools available (e.g., a program for an automatic syllabification), by the researcher's aim, by the possibility to compare results with previous works, etc.

[2] One cannot a priori exclude the existence of some intermediate (maybe not so apparent) levels between the "traditional" ones, see the discussion in Section 5.

[3] Milička (2014) suggested an alternative mathematical model.

[4] Obviously, this special case of the MAL can be applied only under condition of a monotonous relation.

[5] See Altmann (2014) for a bibliography on the MAL.

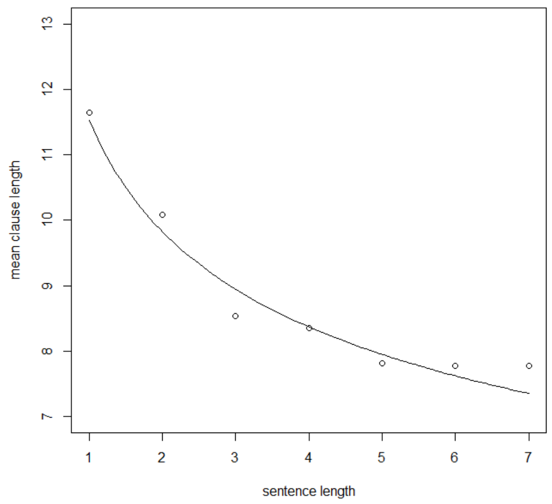

Figure 1. Sentence length (x-axis) and mean clause lengths (y-axis) in a German text (Teupenhayn and Altmann, 1984).

The data can be modelled by the simpler version of the MAL, i.e., by function (2). One obtains parameter values a = 11.571, b = -0.229. The determination coefficient for the model is R^2= 0.9659, indicating thus a very good fit[6].

In general, the interpretation of the parameters of the MAL remains an open question. While Köhler (1984) suggested a very general interpretation, and parameter values resulting from relations between different language units were presented by Cramer (2005b), a connection between the numerical values of the parameters, language levels from which the values arise, and the general theoretical framework (such as a supposed processing capacity, structural information, and similar considerations, see Köhler 1984) is still an unsolved problem[7].

3 Methodology

The MAL predicts that there should be a systematic relation between the size of the clause and the size of its parts. As for the determination of the clause, one can find a (more or less) general agreement among linguists about the character of this unit; e.g., Crystal (2008, p. 78) defines the clause as "a unit of grammatical organization smaller than the sentence, but larger than phrases, words or morphemes". According to the Prague Dependency Treebank annotation[8], which is used for the analysis in this study (see Section 4 for a very brief description, and Lopatková et al., 2009, for more details), clauses "are grammatical units out of which complex sentences are built. A clause typically corresponds to a single proposition expressed by a finite verb and all its arguments and modifiers (unless they constitute clauses of their own).". Regarding the MAL, the clause represents the *construct*.

It is less obvious how to determine parts of the clause which, in accordance to the theoretical background of the MAL (see Section 2), must be defined as its *constituents*. Following both the verb-centric character of dependency syntax traditionally used for Czech and the annotation of the Prague Dependency Treebank, we start with the assumption that the predicate represents the central element of the clause. Thus, the predicate is the highest unit of a hierarchical structure of the clause (see, e.g., Figure 2). Next, all phrases[9] directly dependent on the predicate, i.e. all its arguments and modifiers, are considered *constituents* of the clause (in the sense of the MAL); see Figure 2 where directly dependent phrases are bounded by dashed boxes. Finally, the size of the constituent (i.e., the size of a phrase which is directly dependent on the predicate) is measured by the number of words which the phrase consists of[10].

For an illustration, let us take the clause

My friend saw your sister from Pisa yesterday

depicted in Figure 2.

---

[6] The models in Sections 2 and 5 were fitted to the data by NLREG program.

[7] Another attempt to interpret the MAL – a modification of the ideas from the general approach suggested by Köhler (1984) – can be found in Milička (2014).

[8] http://ufal.mff.cuni.cz/pdt2.5/en/documentation.html# clause

[9] We are aware that Tesnière (2015) used the term *node* (*nœud* in French); however, as the translators of his famous book notice, "[H]he first defines the node to be what modern theories of syntax take to be a phrase/constituent", and "[H]his inconsistent use of the term is a source of confusion" (Tesnière, 2015, Translators' Introduction, p. xlv). We prefer the term phrase in the sense as it is used also by Meľčuk (1988) and Crystal (2008)

[10] We do not claim that this choice of the constituent is the only one possible, or the "right one" for clauses. In our opinion, it is quite probable that there are several "parallel" possibilities, analogous to the chains word – syllable – phoneme and word – morpheme – phoneme. Our approach is the first attempt to investigate the MAL in syntactic dependency structure, and it can be hoped it will be followed by other studies which will open other views.

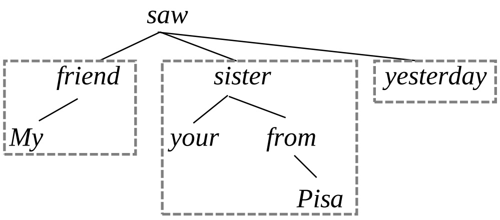

Figure 2. The hierarchical structure of the clause *My friend saw your sister from Pisa yesterday*. The dashed boxes represent phrases which are considered constructs of the clause.

There are three phrases directly dependent on the predicate *saw* (see Figure 2):

(Ph1) *My friend*;
(Ph2) *your sister from Pisa*;
(Ph3) *yesterday*.

Thus, the size of the clause is three. Next, the mean constituent size in the clause is determined as an average of sizes of particular phrases. Specifically, phrase (Ph1) consists of two words, (Ph2) of four words, and (Ph3) of one word. The mean size of the phrase in the clause considered is

$$(2 + 4 + 1) / 3 = 2.33.$$

This procedure is applied to each clause in the corpus (with mean phrase length computed from all clauses with a particular length in the corpus, e.g., we took all phrases which occur in clauses with length one and evaluated their mean length, then all phrases occurring in clauses with length two, etc.).

To sum up,

a) the clause represents the *construct*;
b) the size of the *construct* is determined by the number of phrases which are directly dependent on the predicate of the clause; each phrase represents a *constituent* of the clause;
c) the size of the *constituent* (i.e., of the phrase) is determined by the number of its words.

This approach satisfies the theoretical assumption of the MAL – language units which are in the relation of a construct and a constituent (clause – phrase – word) are used for the analysis.

4 Language material

In this study, dependency trees from the Prague Dependency Treebank 3.0 (Bejček et al., 2013; PDT 3.0 hereafter) were used; specifically, the data annotated on analytical lever (the treebank contains approximately 1.5 million words). Particular clauses from the corpus were determined in accordance with the annotation. Only main clauses were used for modelling because the analytical function "Predicate" is assigned only to the predicate of the main clause in the PDT 3.0[11]. We used tokenized sentences (see Section 3, Figure 2, for an example), with the tokenization from the PDT 3.0 taken without any adaptation. Punctuation is not considered.

Non-projective dependency trees were not filtered out. First, the (non-)projectivity of a dependency tree is irrelevant with respect to the validity of the MAL for the data from a treebank as whole[12]. Clauses consist of phrases regardless of properties of their tree representations. Second, non-projective trees do not present technical problems, as the determination of the predicate and phrases which are directly dependent on the predicate is not affected by the tree (non-)projectivity. Finally, crossings may be not so scarce as it is believed – it seems that they correlate with dependency length (the longer dependency length, the more crossings can be expected, see Ferrer-i-Cancho and Gómez-Rodríguez, 2016). A rejection of non-projective trees could thus lead to an underrepresentation of sentences with longer dependency lengths.

Because of the existence of technical nodes as well as specificities of the annotation in the PDT 3.0, we were forced to rearranged the original annotation to some extent; the whole procedure of the adjustment of the original annotation is described in detail in a technical report which is available online[13].

5 Results

The results – mean lengths of phrases which occur in clauses of particular lengths – are presented in Table 1. Only those clause lengths which occur in the corpus at least ten times, i.e.,

[11] In subordinate clauses, the predicate is not assigned by the analytical function "Predicate" but by a corresponding function of the subordinate clause (e.g., Attribute, Object, Subject).

[12] The validity of the MAL in a subcorpus consisting exclusively of non-projective trees is a different (albeit interesting) question, see a short discussion in Section 6.

[13] http://www.cechradek.cz/publ/2017_macutek_etal_technical _report.zip

up to nine in our case, were analyzed (frequencies of clause lengths measured in the number of phrases in the corpus used can be found in Table 1 as well). Remarks on an irregular behaviour of constituents of long constructs[14] with low frequencies of occurrence can be found, e.g., in Kelih (2010), and in Mačutek and Rovenchak, (2011). The loss of data caused by neglecting longer clauses is minimal. We analyzed 56530 clauses from the corpus (see Section 4), only 18 of them (i.e., approximately 0.03%) consisted of more than nine phrases.

CL	f(CL)	MPL
1	7125	9.47
2	21508	5.04
3	16964	4.00
4	7858	3.51
5	2351	3.25
6	551	2.91
7	118	3.05
8	27	2.85
9	10	3.03

Table 1. Relation between clause length and mean phrase length (CL – clause length, f(CL) – frequency of clauses with the given length in the corpus, MPL – mean phrase length).

The relation can be modelled by the simpler form of the MAL (see Section 2), i.e., by function (2). The parameter values optimized with respect to the goodness of fit (expressed in terms of the determination coefficient) are $a = 8.96$, $b = -0.62$, with $R^2 = 0.9424$. The model fits the data sufficiently well[15] (see Section 2).

The tendency of the mean phrase length to decrease with the increasing clause length can clearly be seen in Figure 3, which depicts also the abovementioned function as the mathematical model for the MAL. We emphasize that the MAL – and all laws in linguistics, and all laws in empirical science in general – is of a stochastic rather than deterministic character, hence some minor local disturbances in the overall decreasing trend are admissible.

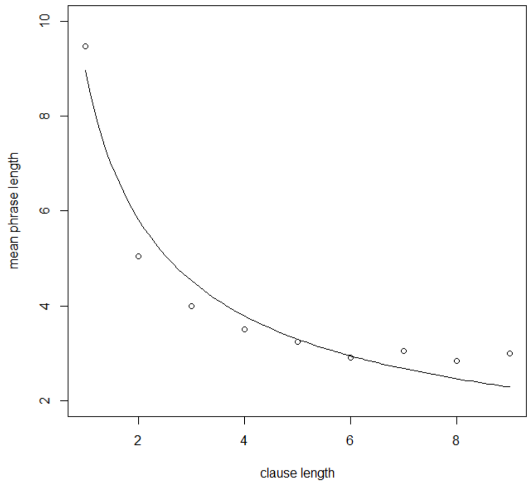

Figure 3. Relation between clause length and mean phrase length (see Table 1), with function (2) fitted to the data.

The result achieved is the first corroboration of the MAL in syntactic dependency structure (some hints towards the validity of the MAL in syntax in general can be found in Köhler, 2012, however, without specifying a wider framework, such as, e.g., dependency grammar in this paper).

6 Conclusion and perspectives

Our paper broadens the scope of the MAL. Based on the analysis of the Czech dependency treebank, it can be said, tentatively at least, that the law is valid also in syntactic dependency structure, with clauses being constructs and phrases (see Section 2, Figure 2) being constituents.

Naturally, further analyses must be postponed until results from several other languages are available. From a theoretical point of view, problems needed to be answered include, e.g., an interpretation of parameters of the mathematical model and relations with other language laws (Köhler, 2005). Another issue waiting to be studied more deeply is the question of non-projective dependency trees. Is the MAL valid for them as well? If yes, do the parameter values

[14] It remains unclear whether the irregular behavior is caused only by low frequencies of long constructs, in which the mean length of constituents then has than a higher variance, or whether there are also other factors at play, which have only a negligible influence on short constructs. Admittedly, if one includes rarely occurring longer constructs, the fit usually becomes worse (which is true also for data considered in this paper).

[15] The "full version" of the MAL, i.e., function (1) from Section 2, achieves a slightly better fit ($R^2 = 0.9970$, with $a = 8.11$, $b = -1.06$, $c = 0.15$), but it has one parameter more, making thus attempts to interpret the parameters more difficult.

differ from the ones typical for corpora in which projective trees prevail?

In more applied fields, parameters of the MAL parameters in dependency structure could perhaps strengthen the arsenal of tools used in authorship attribution, automatic text classification, and similar areas.

The parameters of the MAL in syntactic dependency structure offer themselves to be used in a syntactic language typology (see, e.g., Song, 2001; Whaley, 2010). It would be interesting to take some established typology and to check whether there are some typical parameter values for typologically similar languages. We remind that several attempts to build a language typology based on dependency grammar and on some characteristics of dependency relations appeared in recent years (Liu, 2010, Liu and Li, 2010; Liu and Xu, 2012; Jing and Liu, 2017).

In addition to bringing some results, the paper also opens several questions of theoretical and/or methodological character, some of which can be interesting not only within dependency grammar but also in mathematical modelling of language phenomena in general. We mention some of them in the following paragraphs.

The MAL is usually modelled across neighbouring levels in the language unit hierarchy. It seems that clauses and phrases (as defined in Section 2) are "neighbours" in this sense. The question is which is the next unit when one looks "downwards". We chose word as the constituent of a phrase, but the possibility that we skipped some level(s) cannot be a priori excluded. Will the MAL be valid also for the relation between phrases and "subphrases", i.e., units directly dependent on phrases? If yes, how many levels are there?

Up to our knowledge, there are no published results on the relation between sizes of clauses and words[16]. The paper by Buk and Rovenchak (2008), focusing mainly on the relation between sentence length and clause length (relation between clause length in words and word length in syllables can be reconstructed from the data for a narrow interval of clause size), does not bring any convincing results, it ends with a call for a clarification of the notion of clause. Can the reason be that clauses and words are not neighbours in this sense[17], and that one should consider an intermediate level, such as phrase in this paper?

Nonetheless, the MAL is a good model (in terms of goodness of fit) for the relation between lengths of sentence (in clauses) and clause (in words). The validity of the law was corroborated in eight languages (Czech, English, French, German, Hungarian, Indonesian, Slovak, Swedish), see Köhler (1982), Heups (1983), and Teupenhayn and Altmann (1984). But, as it was mentioned above, clauses and words do not seem to be direct neighbours in the language unit hierarchy. These two facts – the assumed existence of some level(s) between clause and word on the one hand, and the validity of the MAL for the relation between lengths of sentences in clauses and of clauses in words – can be reconciled, e.g., if not one, but two levels (phrases and "subphrases") were omitted. Still another possible explanation is that we analyze parallel nested structures analogous to, e.g., the two chains of units mention in Section 2, one of which consists of words, syllables and phonemes, and the other of words, morphemes and graphemes. Dependency grammar, with its (relatively) clearly defined relations among words in a clause, can be a useful tool for determining "reasonable" (i.e., linguistically interpretable) language units "between" clause and word (if there are any) and for investigating relations among them.

It is our hope that our paper may serve as a stimulus towards future research in the areas of syntactic dependency structure and of relations among language units in general (especially with respect to their sizes and mutual influences).

Acknowledgment

Supported by the VEGA grant no. 2/0047/15 (J. Mačutek) and by the Charles University project Progress 4, Language in the shiftings of time, space, and culture (J. Milička).

[16] Similar discussions were opened by Chen and Liu (2016) on the relation between sizes of word and its constituents (i.e., one level lower than in this paper) in Chinese, and by Sanada (2016) on the relation between sizes of sentence, clause and argument (as defined in Sanada, 2016, pp. 259-260) in Japanese.

[17] According to Köhler (2012, p. 108), "an indirect relationship … is a good enough reason for more variance in the data and a weaker fit".

References

Gabriel Altmann. 1978. Towards a theory of language. In Gabriel Altmann, editor, *Glottometrika 1*, pages 1-25. Brockmeyer, Bochum.

Gabriel Altmann. 1980. Prolegomena to Menzerath's law. In Rüdiger Grotjahn, editor, *Glottometrika 2*, pages 1-10. Brockmeyer, Bochum.

Gabriel Altmann. 1993. Science and linguistics. In Reinhard Köhler and Burghard B. Rieger, editors, *Contributions to Quantitative Linguistics*, pages 3-10. Kluwer, Dordrecht.

Gabriel Altmann. 2014. Bibliography: Menzerath's law. *Glottotheory*, 5(1):121-123.

Eduard Bejček, Eva Hajičová, Jan Hajič, Pavlína Jínová, Václava Kettnerová, Veronika Kolářová, Marie Mikulová, Jiří Mírovský, Anna Nedoluzhko, Jarmila Panevová, Lucie Poláková, Magda Ševčíková, Jan Štěpánek, Šárka Zikánová. 2013. *Prague Dependency Treebank 3.0*. Charles University, Praha.

Solomija Buk and Andrij Rovenchak. 2008. Menzerath-Altmann law for syntactic structures in Ukrainian. *Glottotheory*, 1(1):10-17.

Mario Bunge. 1967. *Scientific Research I, II*. Springer, Berlin.

Heng Chen and Haitao Liu. 2016. How to measure word length in spoken and written Chinese. *Journal of Quantitative Linguistics*, 23(1):5-29.

Irene M. Cramer. 2005a. Das Menzerathsche Gesetz. In Reinhard Köhler, Gabriel Altmann, and Rajmund G. Piotrowski, editors, *Quantitative Linguistics. An International Handbook*, pages 659-688. De Gruyter, Berlin / New York.

Irene M. Cramer. 2005b. The parameters of the Menzerath-Altmann law. *Journal of Quantitative Linguistics*, 12(1):41-52.

David Crystal. 2008. *A Dictionary of Linguistics and Phonetics*. Blackwell, Oxford.

Ramon Ferrer-i-Cancho and Carlos Gómez-Rodríguez. 2016. Crossings as a side effect of dependency lengths. *Complexity*, 21(S2):320-328.

Rainer Gerlach. 1982. Zur Überprüfung des Menzerath'schen Gesetzes im Bereich der Morphologie. In Werner Lehfeldt and Udo Strauss, editors, *Glottometrika 4*, pages 95-102. Brockmeyer, Bochum.

Gabriela Heups. 1983. Untersuchungen zum Verhältnis von Satzlänge zu Clauselänge am Beispiel deutscher Texte verschiedener Textklassen. In Reinhard Köhler and Joachim Boy, editors, *Glottometrika 5*, pages 113-133. Brockmeyer, Bochum.

Yingqi Jing and Haitao Liu. 2017. Dependency distance motifs in 21 Indoeuropean langauges. In Haitao Liu and Junying Liang, editors, *Motifs in Language and Text*, pages 133-150. De Gruyter, Berlin / Boston.

Emmerich Kelih. 2010. Parameter interpretation of Menzerath law: evidence from Serbian. In Peter Grzybek, Emmerich Kelih, and Ján Mačutek, editors, *Text and Language. Structures, Functions, Interrelations, Quantitative Perspectives*, pages 71-79. Praesens, Wien.

Reinhard Köhler. 1982. Das Menzeratsche Gesetz auf Satzebene. In Werner Lehfeldt and Udo Strauss, editors, *Glottometrika 4*, pages 103-113. Brockmeyer, Bochum.

Reinhard Köhler. 1984. Zur Interpretation des Menzerathschen Gesetzes. In Joachim Boy and Reinhard Köhler, editors, *Glottometrika 6*, pages 177-183. Brockmeyer, Bochum.

Reinhard Köhler. 2005. Synergetic linguistics. In Reinhard Köhler, Gabriel Altmann, and Rajmund G. Piotrowski, editors, *Quantitative Linguistics. An International Handbook*, pages 760-774. De Gruyter, Berlin / New York.

Reinhard Köhler. 2012. *Quantitative Syntax Analysis*. De Gruyter, Berlin / Boston.

Haitao Liu. 2010. Dependency direction as a means of word-order typology: a method based on dependency treebanks. *Lingua*, 120:1567-1578.

Haitao Liu and Wenwen Li. 2010. Language clusters based on linguistic complex networks. *Chinese Science Bulletin*, 55(30):3458-3465.

Haitao Liu and Chushan Xu. 2012. Quantitative typological analysis of Romance languages. *Poznań Studies in Contemporary Linguistics*, 48(4):597-625.

Markéta Lopatková, Natalia Klyueva, and Petr Homola. 2009. Annotation of sentence structure: capturing the relationship among clauses in Czech sentences. In *Proceedings of the Third Linguistic Annotation Workshop*, pages 74-81. ACL, Stroudsburg (PA).

Ján Mačutek and Andrij Rovenchak. 2011. Canonical word forms: Menzerath-Altmann law, phonemic length and syllabic length. In Emmerich Kelih, Viktor Levickij, and Yulia Matskulyak, editors, *Issues in Quantitative Linguistics 2*, pages 136-147. RAM-Verlag, Lüdenscheid.

Ján Mačutek and Gejza Wimmer. 2013. Evaluating goodness-of-fit of discrete distribution models in quantitative linguistics. *Journal of Quantitative Linguistics*, 20(3):227-240.

Igor A. Mel'čuk. 1988. *Dependency Syntax: Theory and Practice*. State University of New York Press, Albany (NY).

Paul Menzerath. 1954. *Die Architektonik des deutschen Wortschatzes*. Dümmler, Bonn.

Jiří Milička. 2014. Menzerath's law: the whole is greater than the sum of its parts. *Journal of Quantitative Linguistics*, 21(2):85-99.

Haruko Sanada. 2016. The Menzerath-Altmann law and sentence structure. *Journal of Quantitative Linguistics*, 23(3):256-277.

Jae J. Song. 2001. *Linguistic Typology: Morphology and Syntax*. Routledge, London / New York.

Lucien Tesnière. 2015. *Elements of Structural Syntax*. John Benjamins, Amsterdam.

Regina Teupenhayn and Gabriel Altmann. 1984. Clause length and Menzerath's law. In Joachim Boy and Reinhard Köhler, editors, *Glottometrika 6*, pages 127-138. Brockmeyer, Bochum.

Lindsay Whaley. 2010. Syntactic typology. In Jae J. Song, editor, *The Oxford Handbook of Linguistic Typology*, pages 465-486. Oxford University Press, Oxford.