Acta Universitatis Wratislaviensis
No 3155
Anglica Wratislaviensia XLVII
Wrocław 2009

Agnieszka Kułacka
University of Wrocław

The Necessity of the Menzerath–Altmann Law

In this paper I will discuss the necessity of the Menzerath–Altmann Law to hold. I will present the explanation provided by Köhler, which will be followed by the new insight into the answer of the questions why and in what conditions the law holds.

1. Menzerath–Altmann law

The Menzerath–Altmann law (henceforth MA law) was discovered by A. Grégoire in phonology in 1899. It was later described by Paul Menzerath and fully formulated for the whole language system in the following form (Menzerath 1954 qtd in Hammerl and Sambor 1993: 23): "the longer a language construct the shorter its constituents."[1]

Since then an intense research has taken place in all language levels: phonology, morphology, syntax, discourse and semantics. MA law was verified to hold for Chinese symbols and also in other sciences, such as evolutional biology and genetics. The mathematical model for the law was formulated by Gabriel Altmann in 1980, which was later completed by Agnieszka Kułacka and Jan Mačutek (Kułacka and Mačutek 2007). An attempt to contextualise the coefficients appearing in the formulae was made. In 2008 a new technique of the research was introduced in Kułacka (2008).

In this article I will move closer to answering a profound question why the law holds. Our embankment will be the existing explanation rooted in the theory of the structure of human memory and mechanisms of a language being processed in it. At the very beginning I will provide a description of short term memory since the notion is used by both Köhler and myself. However, before that I will show the reader how MA law works at the syntactic level.

[1] Translated by A.K.

2. MA law in syntax

MA law can be formulated for syntax in the following way: "The longer the clause measured by the number of clauses, the shorter its clauses measured by the number of words." Since both data are discrete, I will use the formula as introduced in Kułacka and Mačutek (2007):

$$y(n) = y(1)\prod_{i=2}^{n}\left(\frac{b}{n-1}+1\right), \tag{1}$$

where $y(n)$ is the sentence consisting of n clauses, $y(1)$ is a simple sentence (a one-clause sentence) and b is a coefficient.

The tables 1 and 2 show the empirical data gathered from first 11 chapters of *The Da Vinci Code* by Dan Brown and first 7 chapters of *Aspects of the Theory of Syntax* by Noam Chomsky, and its theoretical counterparts found by applying formula (1), in which I used the value of coefficient $y(1)$ as obtained from the empirical data. This is the average length of one-clause sentences. To find the value of coefficient b, I used the method of least squares on empirical data. The value can be calculated as the minimum of the following function:

$$f(b) = \sqrt{\sum_{i=1}^{n}(y_{teor}(i) - y_{emp}(i))^2}, \tag{2}$$

where $y_{teor}(i)$ is the length of i-clause sentence as obtained after using formula (1), $y_{emp}(i)$ is the average length of i-clause sentence as obtained from the empirical data. The argument for which the function expressed by formula (2) has its minimum is b. For these calculations I used *Graphic Display Calculator TI-83 Plus* (cf. Kułacka and Mačutek 2007). The procedure to check whether the law holds for a given text applied before the abovementioned procedure of finding the coefficients is fully described in Kułacka (2009).

In the tables the following notation has been used: x is the number of clauses in a sentence, y_{emp} and y_{teor} as described above, n is the number of analysed sentences. I considered only approximately 98% of sentences when finding the coefficients, and, therefore, for the remaining 2% there is no theoretical data found.

Table 1. Data from *The Da Vinci Code* by Dan Brown

x	y_{emp}	n	y_{teor}
1	6.7584	745	6.7584
2	6.1677	465	6.2150
3	6.1976	275	5.9652
4	5.7442	86	5.8053

x	y_{emp}	n	y_{teor}
5	5.5928	28	5.6886
6	5.3627	17	–
7	5.8571	5	–
8	6.4583	3	–
9	6.4444	2	–
Sum		1626	–

The coefficients have the following values: $y(1) = 6.7584$ and $b = -0.0804$

Table 2. Data from *Aspects of the Theory of Syntax* by Noam Chomsky

x	y_{emp}	n	y_{teor}
1	14.6505	103	14.6505
2	12.5000	167	11.8596
3	10.3273	111	10.7300
4	10.1992	64	10.0486
5	10.4000	24	9.5700
6	8.2083	8	9.2054
7	8.6190	3	–
8	9.1250	1	–
9	9.8889	3	–
10	8.7500	2	–
Sum		486	–

The coefficients have the following values: $y(1) = 14.6505$ and $b = -0.1905$

In part 6 of this article I will comment on the differences between the coefficients in literary texts and textbooks.

3. Short term memory

There exist three types of memory: sensory memory, short term memory and long term memory. The information acquired by sensory memory is processed and stored for a limited period of time in short term memory and, if required, stored in long term memory for future use.

The information stored in short term memory becomes erased after having been covered by a new piece of information or with time. The element that survives in working memory is called a memory trace. Let its initial strength be *A*, which depends on the strength of the signal, and let *f* and *g* mean forgetting coefficients having values on [0, 1]. As a result of new information of initial strength *B*, the strength of the first memory trace is lowered to $A \times f^1$. If yet another piece of information is acquired then the strength of the first memory trace is $A \times f^2$ and the strength of the second is $B \times g^1$, etc. It is then assumed that "the strength of the memory trace lowers in geometric progression, dependent on the number of new elements"[2] (Lindsay and Norman 1991: 351).

Forgetting as a result of memory traces being gradually erased with time is the second hypothesis of the cause of short term memory to be limited. As time, *t*, passes the strength of the information, it can be expressed as

$$S(t) = A \times e^{-kt}, \tag{3}$$

where the forgetting function *k* is positive (Lindsay and Norman 1991: 351). As a matter of fact, the limited capacity (and so imposed forgetting) is assumed to be the corroboration of the two causes.

4. Köhler's explanation

Köhler puts forward two assumptions regarding language processing: (1) the coding and decoding of the information is sequential, and (2) the short term memory has limited and constant capacity (cf. Hammerl and Sambor 1993: 45–47), both as having been hypothesised above. The short term memory is responsible for storing the language constituents as long as they are needed for the sequential analysis and synthesis of a language construct, and also for storing the information of the complexity of the construct.

Following the assumptions, it is clear that at each language level, i.e. phonological, morphological and syntactical, there is an upper boundary of the length of these language sequences. Due to that fact, the capacity left for each constituents in a language construction follows the pattern: the more complicated the information structure in terms of the number of its constituents is, the less space for their length is left. The conclusion is "the longer a language construct, the shorter its constituents," i.e. MA law.

5. Capacity constraint comprehension theory

In a *Capacity constraint comprehension theory* storing and processing the information, which takes place in working memory, is due to the process called an activation. The capacity of short term memory, partly[3] used in the process of language

[2] Translated by A.K.

[3] Partly used by other senses and processes.

understanding, can only be limited by an upper boundary of the activation that supports this memory function. Each piece of information has its unique activation level, e.g. the activation level of a lexeme is proportional to its frequency in the mental dictionary and becomes active in the process of coding or decoding of linguistic information. As long as the activation level is above a certain minimal value, the element is kept in short term memory. Otherwise, following the above-mentioned causes, it is erased from it (Just and Carpenter 2002).

Let us now discuss the new insight into the answer to the question why it is necessary for MA law to hold in the context of the *Capacity constraint comprehension theory*.

6. New insight

For better understanding, let us assume that I consider a sentence. The elements of information, which are parts of the sense expressed by the sentence, are the number of words, the complexity of their pronunciation, their morphological and semantical complexity, the complexity of their syntactical structure and the complexity of the syntactical structure of the clauses. For each of them there exists a certain activation level. The number of all activations taking place simultaneously cannot be in excess of the upper boundary, i.e. the capacity of working memory (cf. Just and Carpenter 2002: 135). If then for the activation of syntactical structure of a sentence more capacity of working memory has been reserved, less can be allocated to the complexity for the other elements of the information, e.g. for the complexity of the syntactical structure of the words or their number in clauses. In turn this explains the necessity of MA law to hold: the more complex the sentence, the less complex are its clauses.

In the research that I have conducted on MA law in syntax, the complexity of a sentence was expressed by the number of clauses, and the complexity of a clause was expressed by the number of words used. It is understood that it is the simplified way of a research on MA law; however, the only one that is possible at this state of art. It is conceivable that the law is considered statistical for the very reason of the technique of the research since to measure the real activation level of each of the elements would be too complicated at this stage. Moreover, following the *Capacity constraint comprehension theory*, the processes of activating the elements of information do not take place sequentially, as assumed by Köhler, but simultaneously, which causes additional difficulty during measuring the real activation level of individual elements. The only existing tests examine the capacity of the whole working memory (cf. Just and Carpenter 2002: 136).

In my research I did not take into consideration the differences between the interlocutors. People differ in the capacity of their working memory, which can explain the different values of the coefficients $y(1)$ and b in the formula used.

However, one can formulate the following hypothesis for the types of texts: the value of $y(1)$ found for the novel is less than the value of $y(1)$ found for the textbook, and the opposite is observed for the coefficient *b*. I mainly focussed on the differences between two genres: novels and textbooks,[4] without analysing the differences between the authors. The hypothesis was verified and turned out to be true for all analysed texts. It can easily be explained by Stylistics that, in general, the sentences used in novels are less complex and simple sentences are longer in textbooks.

It was outside my research to discuss the interdependence between the activation levels of individual elements of the information as mentioned earlier. All of the above could be followed up in a future research.

References

Hammerl, R. and J. Sambor. 1993. *O statystycznych prawach językowych.* Warszawa: Zakład Semiotyki Logicznej Uniwersytetu Warszawskiego.

Just, M.A. and P.A. Carpenter. 2002. "A Capacity Theory of Comprehension: Individual Differences Working Memory." In: Polk T.A. and M.C. Seifert. *Cognitive modelling*. London: The MIT Press, 131–177.

Köhler, R. 1984. "Zur Interpretation des Menzerathschen Gesetzes." In: Boy J. and R. Köhler, *Glottometrika* 6, 177–183.

Kułacka, A. 2008. "Badania nad prawem Menzeratha–Altmanna." *LingVaria* 2(6), 167–174.

Kułacka, A. 2009. "Warunki zachodzenia prawa Menzeratha–Altmanna." *LingVaria* 1(7),17–28.

Kułacka, A. and J. Mačutek. 2007. "A discrete formula for the Menzerath–Altmann Law." *Journal of Quantitative Linguistics*, vol. 14 (1), 23–32.

Lindsay, P.H. and D.A. Norman. 1991. *Procesy przetwarzania informacji u człowieka. Wprowadzenie do psychologii.* Warszawa: Wydawnictwo Naukowe PWN.

[4] I analysed the following texts (both Polish and English version): *The Da Vinci Code* by Dan Brown, *Aspects of the Theory of Syntax* by Noam Chomsky, *Hobbit* by J.R.R. Tolkien, *Semantics 2* by John Lyons, *The Outline of Mathematical Logic* by Andrzej Grzegorczyk.