The Menzerath–Altmann Law in Syntax

The aim of my talk is to present the Menzerath-Altmann Law in Polish and English syntax, and to show how meaning of language constructs must be incorporated in syntax analysis i.e. in text disambiguation.

The Menzerath-Altmann law was discovered by Menzerath in 1928 and formulated for fonology. In 1954 Menzerath generalised the statistical fonological relation for text and system, and he stated it in the following way (Menzerath 1954 qt. in Hammerl&Sambor 1993:23):

'the longer a language construct the shorter its constituents'

Altmann in Altmann (1980) described the statistical language relation by means of a functional model after providing a priori reasoning. The relation could be theoretically described by differential equation (1), following the assumptions that the increase in length of a constituent is proportional to its mean length and is inversely proportional to the number of constituents of a given construct; both the constituents and the construct are measured by means of the same unit.

$$\frac{y'}{y} = \frac{b}{x} \tag{1}$$

where y is the mean length of a constituent, y' is its increase in number of units, x is the number of the constituents of the entire construct, and b is a coefficient. After integrating both sides of equation (1), one can obtain the following:

$$y = ax^b \tag{2}$$

where y is the length of a constituent, x is the number of the constituents of the entire construct, and a and b are coefficients.

The Menzerath-Altmann law has been proven to hold in all language levels including language aggregates (see Hřebiček 1995) and semantics (see

Hammerl&Sambor 1993:32-41), and it has also been shown by means of empirical data and its statistical description that it holds in many languages. So far there has been an attempt to prove the law to hold in Polish syntax (see Ruszkowski 1990), however the relation has not been described by a functional model and the nature of the coefficients a and b of equation (2) has not been discussed.

My research on the Menzerath-Altmann law is to be opened with text disambiguation, then I will collect the numerical data and I will examine the degree of approximation of the empirical data to a power function given by equation (2) by the use of statistical tests. I will analyse data of the Polish language and the English language, and I will also consider the data of various literary styles. I expect the law to hold in the research area therefore I will make an attempt at formulating my conjectures concerning the nature of coefficients a and b of equation (2).

References

[1] Altmann, G. 1980. Prolegomena to Menzerath's law. [in:] R. Grotjahn, *Glottometrika 2.* Bochum: Studienverlag Dr. N. Brockmeyer. 1-10.

[2] Hammerl, Rolf i Jadwiga Sambor. 1993. *O statystycznych prawach językowych [About the statistical language laws].* Warszawa: Zakład Semiotyki Logicznej Uniwersytetu Warszawskiego.

[3] Hřebiček, Luděk. 1995. *Text Levels. Language Constructs, Constituents and the Menzerath-Altmann Law.* Tier: Wissenschaftlicher Verlag Trier.

[4] Ruszkowski, Marek.1991. Prawo Menzeratha w badaniach syntaktycznych [The Menzerath–Altmann Law in Syntax]. [in:] *Język Polski LXXI.* Kraków: Komisja Języka Polskiego PAU.187-92.