# MENZERATH'S LAW AND THE CONSTANT FLOW OF LINGUISTIC INFORMATION

AUGUST FENK
GERTRAUD FENK-OCZLON

## MENZERATH'S LAW. GENERALIZATIONS AND SPECIFICATIONS

Menzerath describes two regularities governing the relation between the number of syllables and the number of phonemes in German words. The interpretations given by MENZERATH (1954) are aimed at what nowadays is called "cognitive economy":

> **"I. Die relative Lautzahl nimmt mit steigender Silbenzahl ab, oder mit anderer Formel gesagt: je mehr Silben ein Wort hat, um so (relativ) kürzer (lautärmer) ist es.**
>
> /.../ Es tritt eine **'Sparsamkeitsregel'** in Erscheinung, die sich psychologisch auf eine Ganzheitsregel dieser Art gründet: **je größer das Ganze, um so kleiner die Teile!** Diese Regel /.../ wird aus der Tatsache verständlich, daß das Ganze jeweils 'übersehbar' bleiben muß. Es wäre lohnend, diesen Gedanken weiter zu verfolgen und seine Berechtigung auch auf anderen Gebieten nachzuprüfen." (MENZERATH 1954:100f.)
>
> **"II. Je silbenreicher die Wörter sind, um so geringer wird die Schwankungsbreite der Elementenzahl.** Vielsilbige Wörter sind also in der Lautzahl untereinander ziemlich gleich, während die geringsilbigen Wörter stärker schwanken.
>
> /.../ Die zweite Regel muß sich gleichfalls irgendwie aus der zu I gefolgerten 'Sparsamkeitsregel' ergeben. Die kleinzahlige Ganzheit bleibt offenbar trotz großer Variabilität immer noch überschaubar, während die großzahlige Ganzheit bereits mit dem lautärmsten Wort nahe an die Maximalgrenze heranreicht und darum nicht mehr gut zu vergrößern noch zu komplizieren ist." (MENZERATH 1954:102)

Regularity I states that words composed of a high number of syllables tend to be composed of a "relatively" low number of phonemes. "Relative" to what? Obviously, only the syllable can be the reference point in question. Therefore and according to KÖHLER's (1986:12ff.) reformulation we may transform regularity I into regularity I':

> **I': There is a negative *correlation* between the length of words as measured in syllables, and the length of *syllables* as measured in phonemes.**

*R. Köhler and B.B. Rieger (eds.), Contributions to Quantitative Linguistics, 11–31.*

| syll./word | phon./word | phon./syll. |
|---|---|---|
| $x$ | $y$ | $z$ |
| 1 | 3,8610 | 3,8610 |
| 2 | 5,7348 | 2,8674 |
| 3 | 7,7176 | 2,5725 |
| 4 | 9,6599 | 2,4150 |
| 5 | 11,7848 | 2,3570 |
| 6 | 13,9299 | 2,3217 |
| 7 | 16,6429 | 2,3776 |
| 8 | 18,0000 | 2,2500 |
| 9 | 20,6667 | 2,2963 |

Table 1: An "extract" of MENZERATH's data (1954:96)

For a direct statistical test of regularity I' Menzerath's data-set (MENZERATH 1954:96) was condensed and put into a new matrix. This matrix (Table 1) makes it possible to examine the length of syllables (in phonemes) as a function of the length of words (in syllables). Correlating column $x$ with column $z$ results in a coefficient of $r_{xz} = -0,766$ ($p < 5\%$). Thus, the coefficient of determination ($= RSQ = r^2$) is 0,587. But a look at Figure 1 (Figures 1 - 4 and 6 can be found in the appendix!) and at the lines connecting data points in this diagram reveals the non-linear nature of this function. Grading the number of syllables per word ($x$) logarithmically increases $RSQ$ (0,842). And if we admit quadratic functions, the "correlation" - in the broader sense of the word - is again higher: "$RSQ$" = 0,876.

From Menzerath's general statement - "the bigger the whole, the smaller its parts" (see quotation above) - one might derive a large number of special cases. With just four levels of aggregation (e.g. clause or simple sentence, word, syllable, phoneme, leaving aside complex sentences, compounds, ...) one might construct eleven such special cases, and some of them would prove to be wrong. (For instance: "The bigger the sentence as measured in syllables, the smaller its words as measured in syllables". See the section "Is there a *Positive* Correlation Between the Number of Syllables per Sentence and the Number of Syllables per Word?") But one of these possible deductions - let us call it "regularity III" - is clearly supported by Menzerath's data presented in our Table 1.

**III: The bigger the word as measured in phonemes, the smaller its syllables as measured in phonemes.**

If we know (from the left diagram in Figure 1) that the number of phonemes per syllable ($z$) is closely connected with the number of phonemes per word ($y$) and that $y$ is an almost perfect linear function of the number of syllables per word ($x$) —

$r_{xy} = +0,999!$ - then we have to assume that regularity III holds. The result of the statistical examination: $r_{yz} = -0.747$ ($p < 5\%$), $RSQ = 0,558$. With the number of phonemes per word ($y$) graded logarithmically, $RSQ$ is 0,758. And in the case of the best fitting quadratic function (see the right diagram in Figure 1) "$RSQ$" = 0,855!

Obviously, the *correlational view* offers an adequate operationalisation of Menzerath's law and points out related regularities. This encourages us to go on using correlational methods in the following analyses dealing with relevant language universals.

During the last decade evidence has increased that regularity I is not restricted to German and/or to the word-syllable-relation (e.g. GERLACH 1982, GROTJAHN 1982, KÖHLER 1982, HEUPS 1983, ALTMANN; SCHWIBBE 1989) and that its generalisation in the sense of a linguistic universal is appropriate: "The longer a language construct the shorter its components (constituents)" (ALTMANN 1980:2).

The law in its general form becomes relevant for and applicable to a particular type of cross-linguistic study. Instead of investigating if a regularity found in certain languages can be extended to other instances of language, this type of – in the strict sense of the word – "cross-linguistic" study analyzes the relation between different dimensions (e.g. number of syllables per sentence as a function of number of phonemes per syllable) on the basis of characteristic values - e.g. mean values - within a variety of individual languages. Insofar as all natural languages can be seen as different instances of a system which has to meet certain requirements (concerning distinctivity, economy, ...), differences between languages are non-arbitrary: The variation (within and) between languages on a certain dimension will be connected with the variation of other dimensions, and the constraints and patterns of this concomitance are the central interest of such studies.

## THE GENERALIZED MENZERATH'S LAW EXPLAINING CROSS–LINGUISTIC FUNCTIONS

### A Re–Interpretation of Former Results by means of Menzerath's Law

#### Word Information as a Function of the Number of Syllables per Word

FUCKS (1956) studied the relative frequency of words of different length (1,2,3...$n$ syllables) in 9 different languages. If these word-frequency-data are transformed into bits, the regression between word-information (in bits) and the length of words (in syllables) deviates only very little from a theoretically postulated proportionality-function between the information and the "length" of words as measured in syllables (FENK; FENK 1980). But these deviations do not seem to be accidental. And again the small but systematic deviations can be explained by Menzerath's law: *If longer words tend to be composed of shorter syllables* (MENZERATH 1954) in cross–linguistic comparison as

well, and if we determine the length of words by the number of syllables, the proportionality between information and processing-time has to result in a non–linear, quadratic regression between word-information and "word–length".

Thus, if word information is analyzed as a quadratic function of the number of syllables, the coefficient of determination is higher than in the case of a linear function. (See the "*RSQ*"- values in Figure 2!). And it is scarcely lower than in the case of a cubic regression, which has an even higher degree of freedom in achieving a good fit with real data.

But the linear functions obtained meet better than the quadratic function another requirement of "proportionality", i.e. the requirement of running through the origin of coordinates: As illustrated in Figure 3, the linear functions and their bisector are almost perfect in this respect. Therefore, the advantage of the quadratic regression diminishes if, for theoretical reasons, regressional functions are "forced" through the origin of coordinates. (See Table 2 and Figure 4)

| | linear | quadratic |
|---|---|---|
| 1st function: $\sum_{i=1}^{9}(y^* - y)^2$ | 0,117 | 0,090 |
| 2nd function: $\sum_{i=1}^{9}(x^* - x)^2$ | 0,138 | 0,137 |
| O/O: | | |
| 1st function: $\sum_{i=1}^{9}(y^* - y)^2$ | 0,120 | 0,114 |
| 2nd function: $\sum_{i=1}^{9}(x^* - x)^2$ | 0,146 | 0,137 |

1st function = $y$ as a function of $x$, with $y$-parallel ~~derivations~~ *deviations* minimized
2nd function = $x$ as a function of $y$, with $x$-parallel ~~derivations~~ minimized
O/O: 1st and 2nd function, when "forced" through the origin of coordinates

Table 2: A comparison of linear and quadratic functions regarding their "goodness of fit" with real data $(x^*, y^*)$.

## The Number of Phonemes per Syllable as a Function of the Number of Syllables per Sentence

In an experimental study (FENK-OCZLON 1983) 27 native speakers of 17 Indo-European and 10 Non-Indo-European languages were asked to translate 22 German "kernel-sentences" into their own, typologically different languages and to determine the length of the translations in terms of words and syllables. It was found that the number of syllables varied only within the small range of 7 plus minus 2, and that there is a marked asymmetry in the distribution of languages within this range. (See Table 3)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Dutch | 5.05 | | | | | | | |
| French | 5.3 | | | | | | | |
| Chin. | 5.4 | | | | | | | |
| Czech | 5.4 | | | | | | | |
| Slov. | 5.5 | | | | | | | |
| Hebr. | 5.5 | | | | | $\bar{x} = 6.43$ | | |
| Germ. | 5.5 | | | | | | | |
| Icel. | 5.5 | Bamb. | 6.45 | | | | | |
| Eston. | 5.7 | Turk. | 6.5 | | | | | |
| Russ. | 5.7 | Alban. | 6.5 | | | | | |
| Sbkr. | 5.8 | Port. | 6.6 | | | | | |
| Engl. | 5.8 | Pers. | 6.6 | | | | | |
| Ewon. | 5.8 | Hindi | 6.7 | Ital. | 7.5 | | | |
| Hung. | 5.9 | Pen. | 6.7 | Greek | 7.5 | Anjang 8.2 | | |
| Arab. | 5.9 | Mac. | 6.95 | Span. | 7.9 | Korean 8.2 | | Japan. 10.2 |
| 5 - 5.99 | | 6 - 6.99 | | 7 - 7.99 | | 8 - 8.99 | 9 - 9.99 | 10 - 10.99 |

Table 3: The number of syllables per simple declarative sentence in different languages. (All data, except for Bambara, from FENK-OCZLON 1983)

In order to explain this asymmetry, the number of phonemes per sentence was determined in a later study, and the mean number of phonemes per syllable was correlated with the mean number of syllables per sentence.

The result was a coefficient of $r = -0,77 (p < O,1\%)$. In words:

> **IV: "The higher the mean number of syllables per simple declarative sentence, the lower the mean number of phonemes per syllable."** (FENK-OCZLON; FENK 1985:357)

This system underlying the asymmetric distribution might be regarded as a special case of, or cross-linguistic support for, Menzerath's principle "the bigger the whole, the smaller its parts." And it makes sense with respect to the constant flow of linguistic information: Transmitting *one* proposition with a lower number of syllables demands a higher complexity (and a longer duration) of syllables.

## A Tentative Conclusion

The results presented in the two forgoing sections reveal that the extension of Menzerath's law to cross-linguistic functions is valid. They may be regarded as empirical arguments for the generalized Menzerath's law. In other words: According to the distinction (COOMBS 1984, FENK; VANOUCEK (in press)) between two dimensions of empirical progress - "generality" and "power" - we may talk about a successful attempt to extend the "generality" of Menzerath's law by extending its domain to a new category of empirical instances, i.e. the results of - in the strict sense of the word -

"cross-linguistic" studies. The results give the impression (see Figure 5 in Appendix), that Menzerath's law is apt to explain

- the deviations from a random dispersion of syllables per simple declarative sentence in different languages (Table 3)
- the deviations from the strict proportionality between word-length in syllables and word-information. (The functions obtained indicate, moreover, that relevant theories might achieve higher precision or "power" when using measures and descriptions of information theory.)

**Further Deductions and their Examination**

**Syllables per Word as a Function of Words per Sentence**

If, in cross-linguistic comparison, the mean number of syllables per simple declarative sentence is "constant" (i.e.: if it varies only within a small range), then in languages using more words for forming a sentence the number of syllables per word has to be lower.

Again this hypothesis derived from the principle of a constant flow of information coincides with the generalized form of Menzerath's law. It says:

> **V: Computed across different languages there is a negative correlation between the "size" of sentences as measured in words and the "size" of words as measured in syllables.**

This prediction was examined using our data displayed in Table 4. (Data in columns U and V originate from FENK-OCZLON (1983), those under Y from FENK-OCZLON; FENK (1985). Only Bambara was investigated and included later on.) The result: $r = -0,692$ ($n = 29$; highly significant, $p < 0,1$ %). The coefficient of determination is 0,479 in the case of the linear function and 0,504 in the case of a quadratic function. (See Figure 6 in the appendix!)

| | language | words/sent. U | syll./sent. V | phon./sent. W | syll./word X | phon./syll. Y | phon./word Z |
|---|---|---|---|---|---|---|---|
| 1 | Arabic | 2,455 | 5,955 | 15,273 | 2,4257 | 2,5647 | 6,2212 |
| 2 | Russian | 2,545 | 5,682 | 13,545 | 2,2326 | 2,3838 | 5,3222 |
| 3 | Turkish | 2,591 | 6,455 | 14,636 | 2,4913 | 2,2674 | 5,6488 |
| 4 | Estonian | 2,591 | 5,681 | 13,591 | 2,1926 | 2,3924 | 5,2455 |
| 5 | Czech | 2,773 | 5,364 | 12,818 | 1,9344 | 2,3896 | 4,6224 |
| 6 | Hebrew | 2,864 | 5,455 | | 1,9047 | | |
| 7 | Slovenian | 2,864 | 5,500 | 12,455 | 1,9204 | 2,2645 | 4,3488 |
| 8 | Serbo-Croatian | 2,955 | 5,772 | 13,500 | 1,9533 | 2,3389 | 4,5685 |
| 9 | Islandic | 3,045 | 5,500 | 15,682 | 1,8062 | 2,8513 | 5,1501 |
| 10 | Korean | 3,090 | 8,182 | 18,909 | 2,6479 | 2,3110 | 6,1194 |
| 11 | Macedonian | 3,182 | 6,955 | 15,318 | 2,1857 | 2,2024 | 4,8140 |
| 12 | Persian | 3,364 | 6,636 | 15,955 | 1,9727 | 2,4043 | 4,7429 |
| 13 | Ewondo | 3,409 | 5,773 | 14,273 | 1,6935 | 2,4724 | 4,1869 |
| 14 | Hungarian | 3,545 | 5,909 | 13,409 | 1,6669 | 2,2693 | 3,7825 |
| 15 | Greek | 3,682 | 7,545 | 15,182 | 2,0492 | 2,0122 | 4,1233 |
| 16 | Albanian | 3,727 | 6,545 | 15,227 | 1,7561 | 2,3265 | 4,0856 |
| 17 | Bambara | 3,777 | 6,455 | 13,636 | 1,7090 | 2,1125 | 3,6103 |
| 18 | Italian | 3,909 | 7,500 | 15,909 | 1,9186 | 2,1212 | 4,0698 |
| 19 | German | 3,955 | 5,500 | 15,636 | 1,3906 | 2,8429 | 3,9535 |
| 20 | Dutch | 4,000 | 5,045 | 15,000 | 1,2613 | 2,9732 | 3,7500 |
| 21 | Portuguese | 4,000 | 6,636 | 14,591 | 1,6590 | 2,1988 | 3,6478 |
| 22 | Spanish | 4,182 | 7,955 | 16,636 | 1,9022 | 2,0913 | 3,9780 |
| 23 | Hindi | 4,182 | 6,773 | 15,409 | 1,6196 | 2,2751 | 3,6846 |
| 24 | Annang | 4,182 | 8,227 | 15,818 | 1,9672 | 1,9227 | 3,7824 |
| 25 | French | 4,227 | 5,318 | 13,136 | 1,2581 | 2,4701 | 3,1076 |
| 26 | Pandshabi | 4,318 | 6,773 | 15,955 | 1,5686 | 2,3557 | 3,6950 |
| 27 | English | 4,364 | 5,772 | 15,500 | 1,3226 | 2,6854 | 3,5518 |
| 28 | Japanese | 5,227 | 10,227 | 19,182 | 1,9566 | 1,8756 | 3,6698 |
| 29 | Chinese | 5,409 | 5,409 | 14,727 | 1,0000 | 2,7227 | 2,7227 |
| ∅ | | 3,600 | 6,431 | 15,032 | 1,8402 | 2,3606 | 4,2931 |

Table 4: The mean number of "elements" (word, syllables, phonemes) in 22 simple declarative sentences.

**Is There a *Positive* Correlation Between the Number of Syllables per Sentence and the Number of Syllables per Word?**

Differentiating between only four levels of aggregation (A sentence, B word, C syllable, D phoneme) permits, as already mentioned, eleven derivations from the principle "the bigger the whole, the smaller its parts." Four out of these special cases form a group which includes regularity I' and three other regularities which share the form of I' insofar as there is one member in a "middle position": This "part" is the component and measure of the bigger construct *and* is itself measured by the number of its components.

| | |
|---|---|
| (a) A - *B* - C | The bigger the sentence in words, the smaller the word in syllables. |
| (b) B - *C* - D | The bigger the word in syllables, the smaller the syllable in phonemes. |
| (c) A - *B* - - - D | The bigger the sentence in words, the smaller the word in phonemes. |
| (d) A- - - *C* - D | The bigger the sentence in syllables, the smaller the syllable in phonemes. |

(a) and (d) are a direct consequence of the constant information flow and have proved to hold in cross-linguistic comparison. Direct statistical support for (a) is reported in the foregoing section, and for (d), in the section "The Number of Phonemes per Syllable as a Function of the Number of Syllables per Sentence".
In (b) and (d) the syllable takes the "middle-position", and we can link them together in a "syllogism":

- Premise 1: Languages with a less complex syllable structure (fewer phonemes per syllable) tend to produce *sentences* with a higher number of syllables.
- Premise 2: Languages with a less complex syllable structure (fewer phonemes per syllable) tend to produce *words* with a higher number of syllables.
- Inference: In cross-linguistic comparison we will find a positive correlation between the number of syllables per sentence and the number of syllables per word forming these sentences.

This inference in other words:

> **VI: The bigger the sentence as measured in syllables, the bigger the word as measured in syllables.**

The result of a statistical examination of this conclusion by means of the data in Table 4:

$$r = +0,376\ (n = 29; p < 5\%),\ RSQ = 0,141$$

(with a logarithmic gradation $RSQ$ = 0,161, and with a quadratic function "$RSQ$" = 0,201).

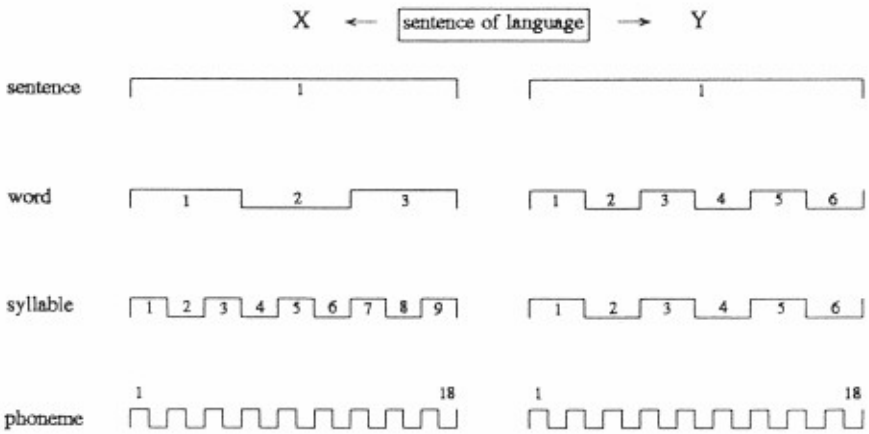


Fig. 7: A graphical illustration of two premises and the relevant conclusion (see text):

| | |
|---|---|
| Premise 1: | The language with phonemically poor syllables (language X) produces sentences with a higher number of syllables |
| Premise 2: | The language with phonemically poor syllables (language X) produces words with a higher number of syllables |
| Conclusion: | A higher number of syllables per sentence coincidence (in language X) with a higher number of syllables per word |

Although both premises coincide with the principle "the bigger the whole, the smaller its parts", the inference drawn contradicts this principle. (Obviously, the application of this principle is dangerous in propositions of a different type than exemplified by our examples (a) to (d). In the context of "Arens's law", a similar paradox is discussed by ALTMANN; SCHWIBBE 1989:46-48). A simple model illustrating the compatibility of regularity I' (premise 2) and our conclusion is presented in Figure 7.

Premise 2, i.e., the hypothesis that regularity I' will be valid in cross-linguistic comparison, is - indirectly - supported by the findings reported. The results of a *direct* examination are: The linear correlation between the number of phonemes per syllable and the number of syllables per word is

$r = -0,452 (p < 1\ \%)$, $RSQ$ = 0,204 (In the case of a quadratic regression "$RSQ$" is 0,258)

Thus we may state an additional regularity:

**VII: Computed across different languages, there is a negative correlation between the "size" of words in syllables and the "size" of syllables in phonemes.**

As already mentioned, premise 1 is a direct consequence of a constant flow of information. One might even argue that premise 2 is the consequence of premise 1 under the presuppositon that the number of words per sentence varies only within a small range and/or independently of the number of syllables per sentence. (To be more precise: If we replace premise 2 in our old syllogism by this presupposition, the conclusion of this new syllogism is identical with premise 2 in the old syllogism.) If, for example, the mean number of syllables is 6 in language *x* and 9 in language *y*, and the mean number of words is 3 in both languages, then the mean number of syllables per word is 2 in the case of *x* und 3 in the case of *y*.

## MENZERATH'S LAW EXPLAINED BY PERCEPTIVE AND COGNITIVE MECHANISMS

If the aim is to find an explanation for Menzerath's law itself, mechanisms of perception and cognition should be considered.

In this context, KÖHLER (1989) argues that the processor(s) involved might have less capacity available for the processing of a construct's single components as the "structural information" (concerning the interaction between the components of a more complex construct) increases. Thus, in order to meet such constraints, more complex constructs might tend to be composed of shorter components. This interpretation of Menzerath's law raises the question of whether - or at which level of complexity - the integration of an element into a supersign is not sufficient to reduce the information of single elements to the required extent. A similar question is raised by Menzerath's hint that it is the function of regularity I and II to keep the bigger construct "übersehbar", "überschaubar": What is it that makes a temporal sequence of elements "comprehensible at a glance", and what is it that constrains the length of a series which is "comprehensible at a glance"?

The "psychological present" is said to have a maximal duration of 1,5 - 3 seconds (FRAISSE 1957/1985:95, BADDELY; THOMSON; BUCHANAN 1975:575, PÖPPEL 1985) and to comprise up to 7 (plus minus two) elements. Allen in his literature review:

> "When we hear a sequence of pulses that is neither too rapid nor too slow we hear it as rhythmic /.../ As long as the minimum time between pulses is greater than about 0.1 s, so that successiveness and order are perceivable, and the maximum is less than about 3,0 s, beyond which groupings do not form, we will impose some rhythmic structure on the sequence. With regular sequences of stimuli, such as a sequence of nearly identically spaced nearly identical clicks, the structures usually perceived are simple groupings of from two to six successive stimuli per group, with faster rates of succession giving more stimuli per group /.../" (ALLEN 1975:76)

The "psychological present" or the immediate memory span may be operative primarily at the sentence processing level and only indirectly at the word and syllable level. Relatively high complexity (allowing high informational content and demanding longer duration) of units at level *n* will result in relatively low complexity, low informational content and short duration at level *n* - 1, if information-related and/or time-related limits

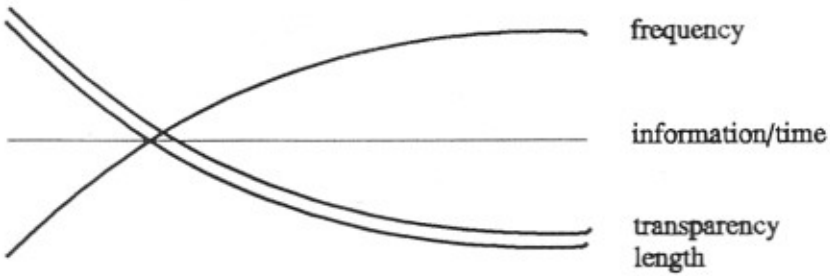


Fig. 8: An increasing frequency of a (super)sign goes hand in hand with an increasing "erosion" of this (super)sign: It shortens and becomes less transparent. Negative effects of the erosion are counterbalanced by higher familiarity.

of our cognitive capacity are efficient at level *n*.

| | words/sentence | syll./sentence | phon./sentence |
|---|---|---|---|
| syllables/word | -0,692 0,1 % | +0,371 5,0 % | |
| phonemes/syllable | | -0,757 0,1 % | |
| phonemes/word | | | +0,103 n.s. |

Table 5: Correlations between the length of sentences (in words, in syllables, in phonemes) and the length of their components.

Such mechanisms should be effective in diachronic changes as well: With increasing token frequency a more complex composition (e.g. a compound) becomes more familiar and, by "erosion", a less transparent but shorter unit, which now offers itself as a component of new compositions (FENK; FENK-OCZLON 1987) As illustrated in Figure 8, the information transmitted and to be processed per unit of time remains constant because the loss in the duration and in the transparency of a sign is compensated for by higher familiarity.

The fact that in cross-linguistic comparison the number of syllables per simple sentence was found to be located in the area of 5 - 9 syllables (see Table 3), agrees with our immediate memory span comprising about 5 - 7 units. And the location in this area corresponds to time-related limits, which might be operative at the level of syllable perception (and production): 200 - 300 milliseconds seems to be the duration necessary for auditive pattern recognition (MASSARO 1975) and for producing the right-ear advantage in dichotic-listening experiments:

"Laurain King and I found that the briefest duration that yielded a right-ear superiority was about 200 milliseconds, or about the duration of an average spoken syllable: a consonant and a vowel. That size of unit seems to be necessary, although not always sufficient, for asymmetrical processing, and it supports the notion that the syllable is a basic unit in speech". (KIMURA 1976:247)

If the duration of a simple sentence coincides with our "psychological present" (c. 2 seconds) and if the minimum duration of a syllable is estimated at c. 200 milliseconds, then the sentence comprises 10 syllables in a "pure CV-language" (see Japanese in Table 2) and a lower number of syllables in the case of more complex syllables (CVC, CCVC, CCVCC, ...), proportionate to the longer duration of these more complex syllables. In this respect, at least, there seems to be nothing magical in the "magical number seven".

The upper limit (2 - 3 sec. per clause or simple sentence) and the lower limit (200 - 300 millisec. per syllable) are operative in the rhythmic pattern organisation, and they might be operative like set points in the self-regulation of language systems, constraining for instance the typological differentiation of languages with regard to morphosyntactic structure and complexity of syllables.

## FINAL CONCLUSIONS

Four *cross-linguistic* "laws" have been presented in the sections above:

IV The more syllables per sentence, the fewer phonemes per syllable.
$r = -0{,}77$ s.

V The more words per sentence, the fewer syllables per word.
$r = -0{,}69$ s.

VI The more syllables per sentence, the *more* syllables per word.
$r = +0{,}38$ s.

VII The more syllables per word, the fewer phonemes per syllable.
$r = -0{,}45$ s.

Only IV was already stated in an earlier study (FENK-OCZLON; FENK 1985). Together with VII - i.e. the cross-linguistic version of Menzerath's regularity I - it forms the premises of a syllogism with VI as the inference drawn. All of these regularities correspond - more (see regularity V!) or less directly - with the principle of a constant flow of linguistic information, and in the interpretation suggested this principle plays the role of a "covering law".

We first discussed Menzerath's law in the role of an *explanans*, and our tentative conclusion (at the end of the section "A Re-Interpretation of Former Results by means of Menzerath's Law") was supported by further results reported in the section "Further Deductions and their Examination". We then (in the foregoing section) discussed Menzerath's law in the role of the *explanandum*, i.e. as the object of the attempted explanation. In both cases arguments seem to boil down to the view that Menzerath's law serves the "constant" and "economic" flow of linguistic information, avoiding an

overcharge as well as a waste of cognitive resources. The limitation of our information processing capacity which necessitates this economic use of cognitive resources is likely to be the most general principle in the continuum of laws illustrated in Figure 5.

Menzerath's "Sparsamkeitsregel" should probably be operationalized in terms of information theory. However, limitations of man's information processing capacity are "universal", and therefore Menzerath's "Sparsamkeitsregel" is effective in each single language and is responsible for the simple mathematical functions found in cross-linguistic comparison. In other words: The fact that typologically very different languages form such functions is a strong indication of the effectiveness of constraints calling for economy principles in the processing of (linguistic) information.

# APPENDIX

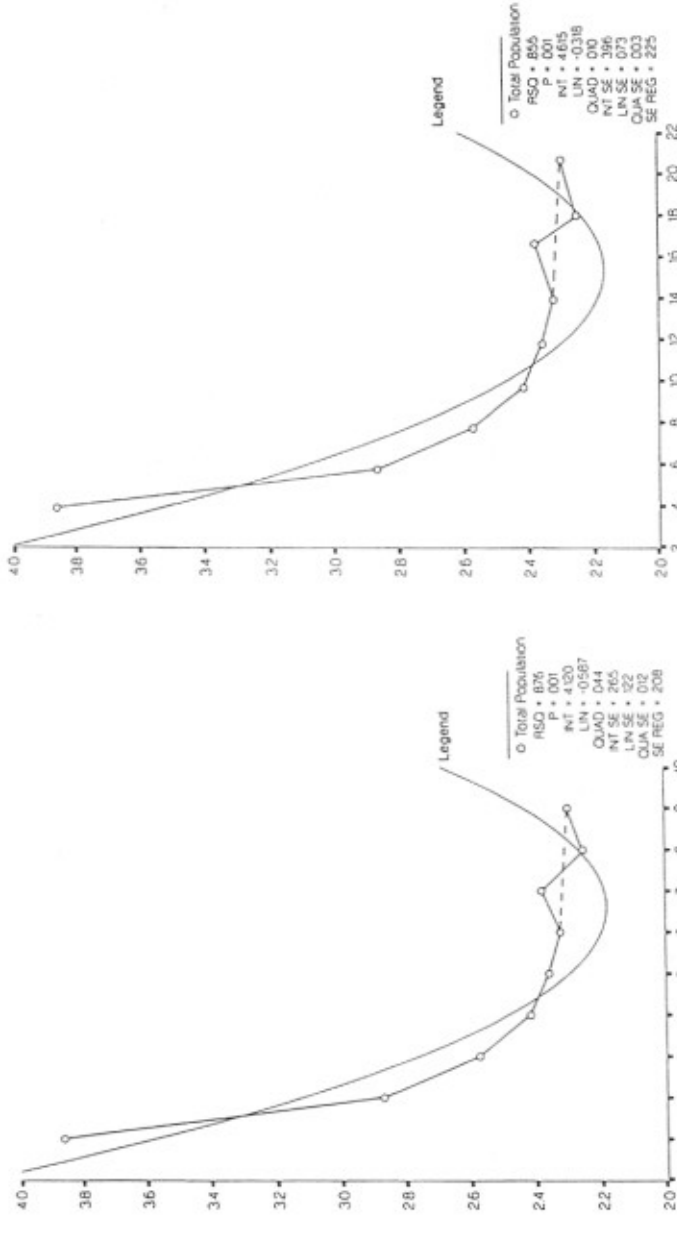


Fig. 1: The number of phonemes per syllable as a function of the number of syllables per word (left diagram) and the number of phonemes per word (right diagram).

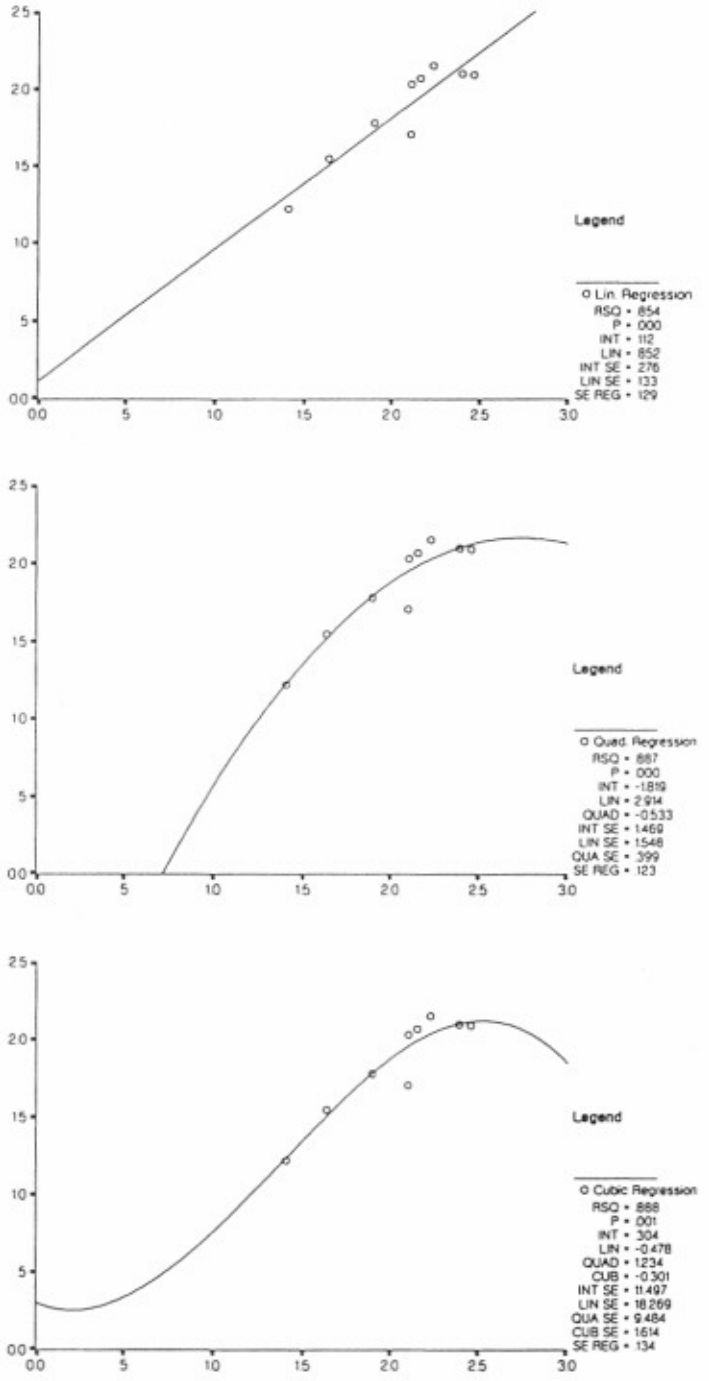


Fig.2: Word information in bits (y-axis) as a linear, a quadratic and a cubic function of word length in syllables (x-axis).

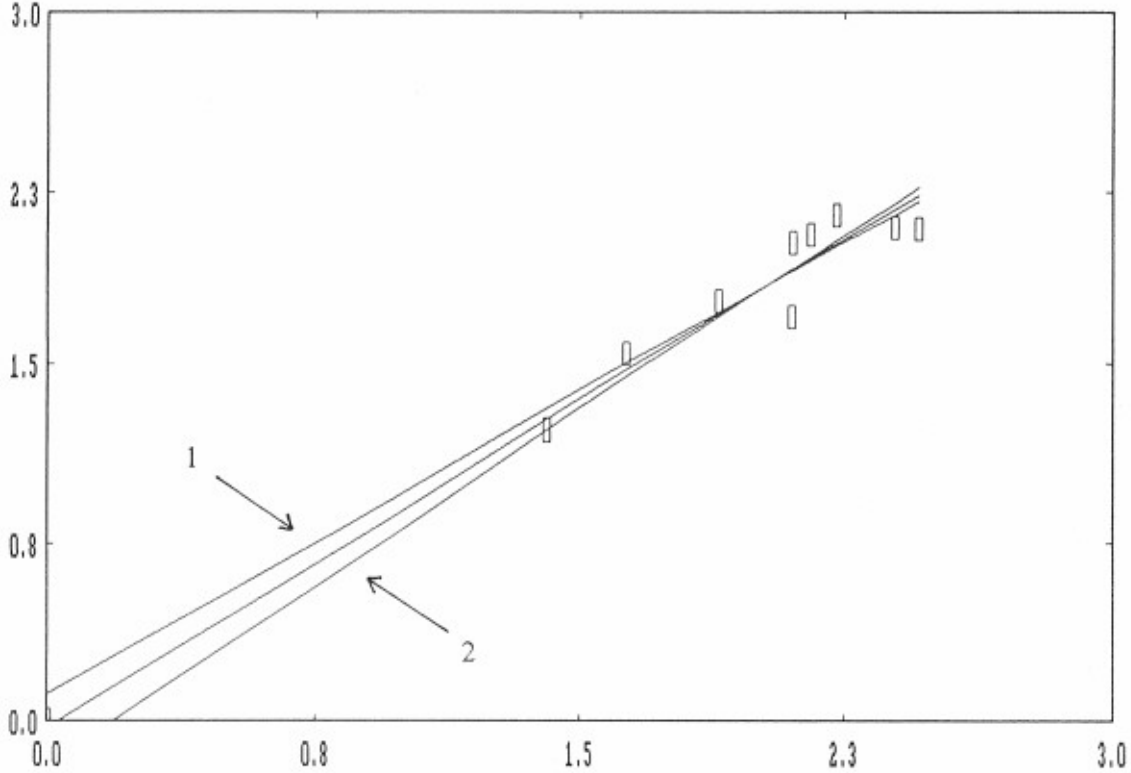


Fig.3: Word information in bits (y-axis) as a linear ~~and as a quadratic~~ function of word length in syllables (1), x as a function of y (2) and the bisection of these two functions.

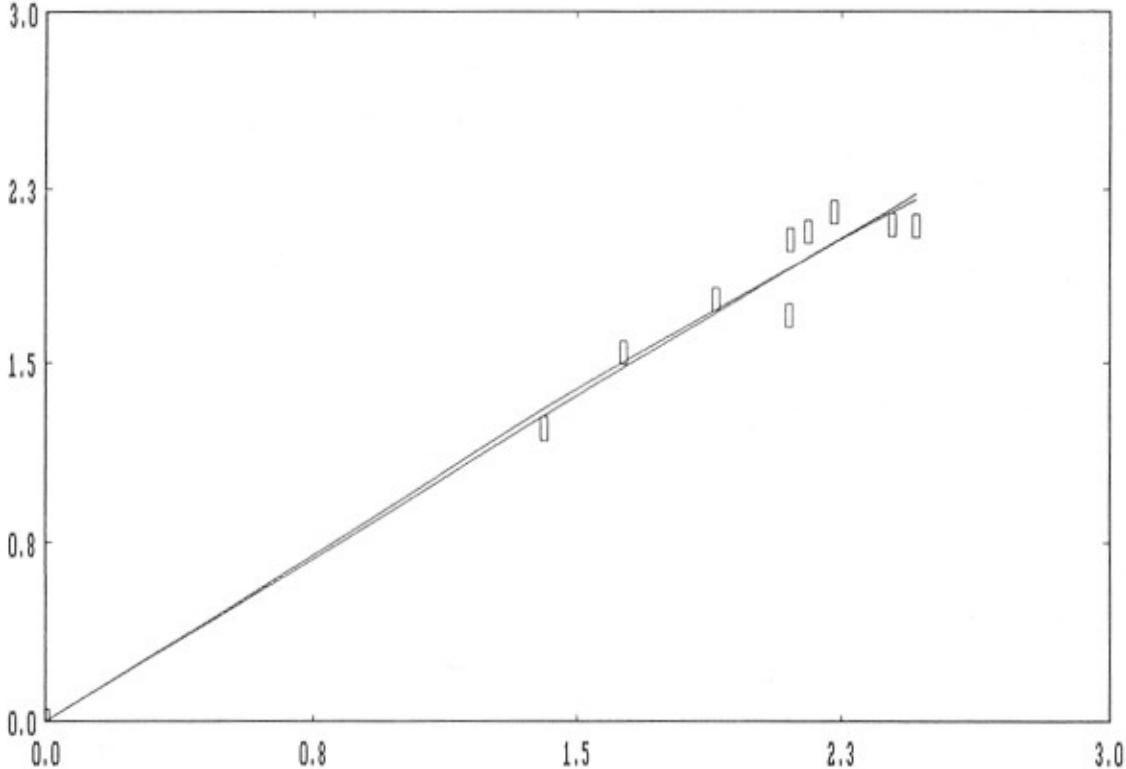


Fig.4: Word information in bits (y-axis) as a linear and as a quadratic function of word length in syllables (x-axis), with both functions "forced" through the origin of coordinates.

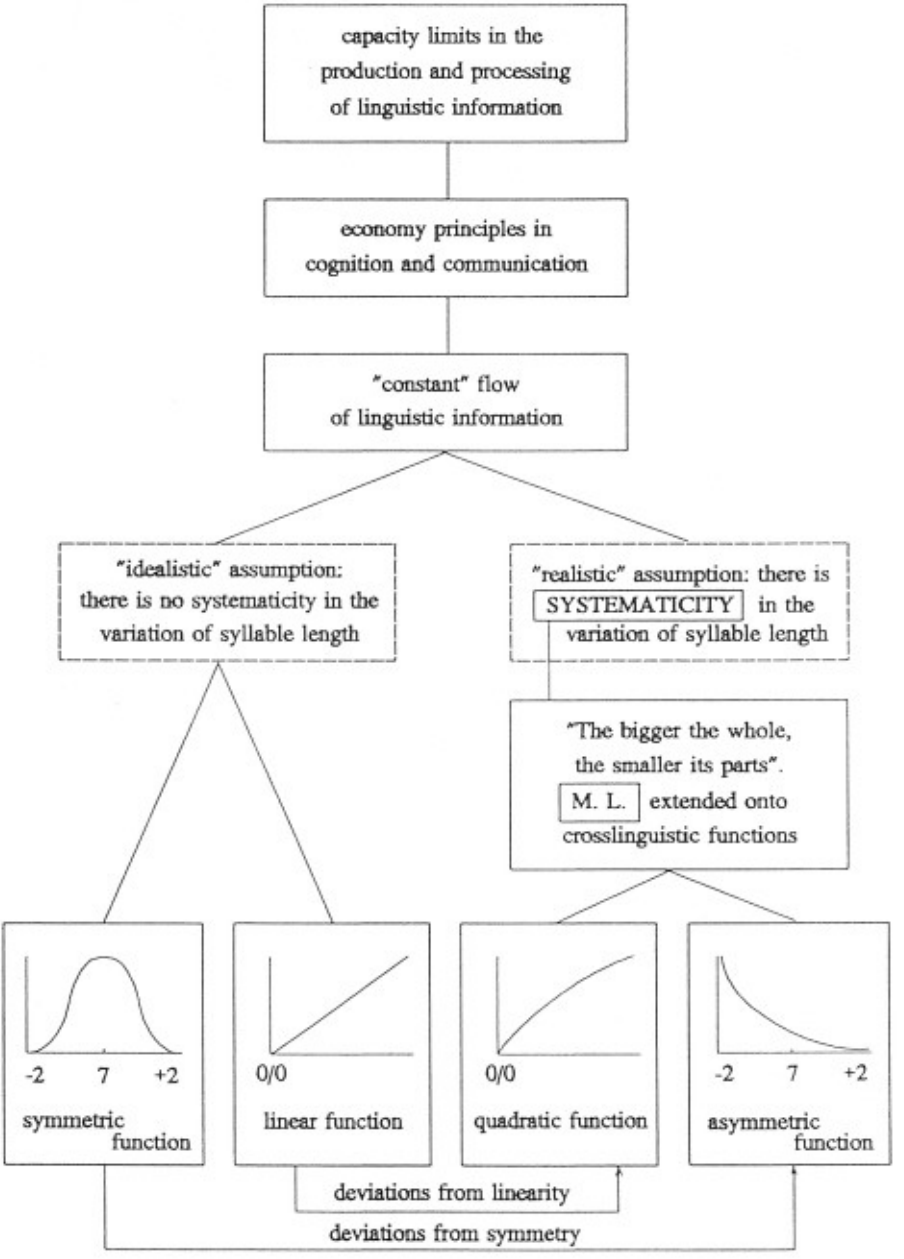


Fig.5: A hierarchy of "laws". (ML = Menzarath's law).

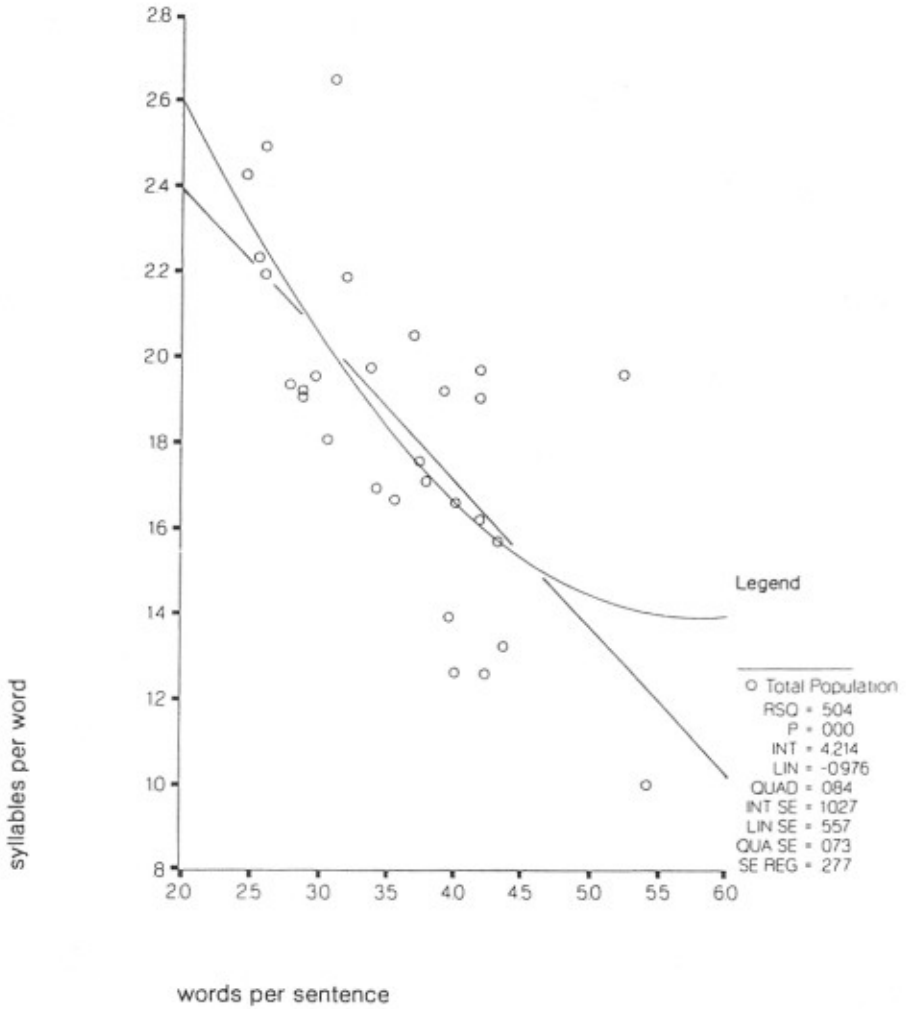


Fig.6: The number of syllables per word as (a linear and) a quadratic function of the number of words per sentence.

## REFERENCES

**Allen, G.D. (1975):** Speech rhythm: its relation to performance universals and articulatory timing. In: *Journal of Phonetics* 3, 75 - 86

**Altmann, G. (1980):** Prolegomena to Menzerath's Law. In Grothjahn, R. (Hrsg.): *Glottometrika* 2. Bochum: Brockmeyer, 1 - 10.

**Altmann, G.; Schwibbe, M.H. (1989):** *Das Menzerathsche Gesetz in informationsverarbeitenden Systemen.* Hildesheim: Olms Verlag

**Baddely, A.D.; Thomson, N.; Buchanan, M. (1975):** Word length and the structure of short-term memory. In: *Journal of verbal learning and verbal behavior* 14, 575 -589

**Coombs, C.H. (1984):** Theory and experiment in psychology. In Pawlik, K. (Hrsg.): *Fortschritte der Experimentalpsychologie.* Berlin-Heidelberg: Springer, 20 - 30

**Fenk, A.; Fenk-Oczlon, G. (1980):** Konstanz im Kurzzeitgedächtnis - Konstanz im sprachlichen Informationsfluß? In: *Zeitschrift für experimentelle und angewandte Psychologie* 27, 400 - 414

**Fenk, A.; Fenk-Oczlon, G. (1987):** Semiotics and the logic of empirical research in naturalness theory. Paper presented at XVIth Congress of linguists. Abstracts: 452. Berlin

**Fenk, A.; Vanoucek, J. (in press):** Zur Messung prognostischer Leistung. In: *Zeitschrift für experimentelle und angewandte Psychologie*

**Fenk-Oczlon, G. (1983):** *Bedeutungseinheiten und sprachliche Segmentierung. Eine sprachvergleichende Untersuchung über kognitive Determinanten der Kernsatzlänge.* Tübingen: Narr

**Fenk-Oczlon, G.; Fenk, A. (1985):** The mean length of propositions is seven plus minus two syllables - but the position of languages within this range is not accidental. In d'Ydewalle, G. (ed.): *Cognition, information processing, and motivation.* Amsterdam: Elsevier Science Publishers, 355 - 359

**Fraisse, P. (1985):** *Psychologie der Zeit.* München: Ernst Reinhardt. (Original edition: *Psychologie du temps.* Paris: Presses Universitaires de France 1957)

**Fucks, W. (1956/1964):** Die mathematischen Gesetze der Bildung von Sprachelementen aus ihren Bestandteilen. In: *Nachrichtentechnische Fachberichte*, Beiheft der NTZ, Bd. 3, 7 - 21

**Gerlach, R. (1982):** Zur Überprüfung des Menzerath'schen Gesetzes im Bereich der Morphologie. In Lehfeldt, W.; Strauss, U. (Hrsg.): *Glottometrika* 4. Bochum: Brockmeyer, 95 -102

**Grotjahn, R. (1982):** Ein statistisches Modell für die Verteilung der Wortlänge. In: *Zeitschrift für Sprachwissenschaft*, 44 - 75

**Heups, G. (1983):** Untersuchungen zum Verhältnis von Satzlänge zu Clauselänge am Beispiel deutscher Texte verschiedener Textklassen. In Köhler, R.; Boy, J. (Hrsg.): *Glottometrika* 5. Bochum: Brockmeyer, 113 - 133

**Kimura, D. (1976):** The asymmetry of the human brain. In: *Recent progress in perception*. Readings from Scientific American. San Francisco: Freeman and Company

**Köhler, R. (1982):** Das Menzerathsche Gesetz auf Satzebene. In Lehfeldt, W.; Strauss, U. (Hrsg.): *Glottometrika* 4. Bochum: Brockmeyer, 103 - 113

**Köhler, R. (1986):** *Zur linguistischen Synergetik: Struktur und Dynamik der Lexik.* Bochum: Brockmeyer

**Köhler,R. (1989):** Das Menzerathsche Gesetz als Resultat des Sprachverarbeitungsmechanismus. In: Altmann; Schwibbe (1989).

**Massaro, D.W. (1975):** *Experimental psychology and information processing.* Chicago: Rand McNally College Publishing Company

**Menzerath, P. (1954):** Die Architektonik des deutschen Wortschatzes. In: *Phonetische Studien*, Heft 3. Bonn: Ferd. Dümmlers Verlag

**Pöppel,E. (1985):** *Grenzen des Bewußtseins.* Stuttgart: Deutsche Verlags-Anstalt