in: V. Solovyev, V. Polyakov (eds.) Text Processing and Cognitive Technologies. Paper Collection No 20. XIII th International Conference "Cognitive Modeling in Linguistics"*, Corfu, September 22-29, 2011. Kazan: Kazan State University Press. pp.161-164*

ABDUCTIVE INFERENCES IN SYSTEMIC TYPOLOGY[1]

August Fenk[2]

Gertraud Fenk-Oczlon[3]

KEYWORDS

Abductive Inferences, Systemic Typology, Statistical Laws, Cross-linguistic Correlations

In this paper we do not present any new empirical results but illustrate, instead, applications of a strategy that allows generating new hypotheses. The focus will be on what is called "law-abduction" in a recent taxonomy [1]. For decades, theory of science was concerned with the logical foundation of explanations and the like; the way empirical assumptions come into world was outside of its focus. Even in books dedicated to "the logic of scientific discovery" [2] and "the growth of scientific knowledge" [3], the term *abduction* does not occur, and the "founding father" Peirce [4] appears, if at all, in footnotes or as one of the proponents of an untenable *subjective theory* [2:361]. Both induction and abduction are, in contrast to deduction, ampliative and uncertain; thus, the conclusion may be subject to further testing. But abduction is a principle in its own right and an appropriate strategy to reveal new causes or explanations [1]. According to Reichenbach [5:400f], a "system of concatenated inductions /.../ is better than any single induction" and "the best posit we know concerning the future." Such a system is, to our opinion, a presupposition for a successful use of abductions to generate new and lawlike hypotheses

Here we describe the application of such a strategy in systemic typology. The program of what we call systemic typology [6] was already formulated by Georg von der Gabelentz [7], in very modern ways and long before the development of systems theory and of those statistical methods allowing a realization of that program: He views language as a "system" the parts of which "organically" interact; any change in one of these parts – morphology,

[2] Alpen-Adria University of Klagenfurt, Austria, august.fenk@uni-klu.ac.at
[3] Alpen-Adria University of Klagenfurt, Austria, gertraud.fenk@uni-klu.ac.at

syntax, sound structure – would change the system as a whole; in the future it should become possible to predict, from the knowledge of a certain property in a certain language, other properties of that language.

The strategy of abductive inferences allows to proceed from already given *cross-linguistic correlations* (one data-pair from each language) to plausible hypotheses regarding further regularities: Two or more of the given regularities ("statistical laws") are linked together as the premises of a syllogism, and the inference drawn is a lawlike assumption that can in turn be tested in the form of a cross-linguistic correlation. (In cases where the correlations taken as "premises" explain a sufficient percentage of the total variation, the respective inference may be not only a plausible but even a cogent conclusion.) For example [8:18]:

A: the more syllables per sentence, *the fewer phonemes per syllable*
B: *the fewer phonemes per syllables*, the more syllables per word
C: the more syllables per sentence, the more syllables per word

One of several possible paraphrases of that syllogism:

A: the fewer syllables per sentence, *the more phonemes per syllable*
B: *the more phonemes per syllable*, the fewer syllables per word
C: the fewer syllables per sentence, the fewer syllables per word

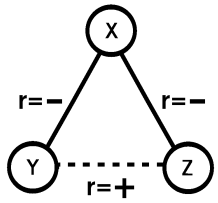

Figure 1: The figure of arguments in diagrammatic representation.
x = n of phonemes per syllable; y = n of syllables per sentence; z = n of syllables per word

Figure 1 illustrates the general line of arguments: A correlation of x with both y and z indicates a possible correlation between y and z. More specifically, it illustrates the role of the positive and negative signs in our correlations: In cases of equal signs in the already given correlations we rather assume a positive sign in the "missing link", otherwise a negative sign. The above inference turned out to be significant. It follows a pattern called "law-abductions"; these abductions "correspond to what Aristotle has called the mind's power of *hitting upon the middle term* of a syllogism" [1]. In our above example the number of phonemes per

syllable (x) is that middle term. But in contrast to the examples presented in [1], all our propositions – premises as well as conclusions - are statistical correlations, i.e., statistical instead of universal laws. And while abduction is, since the times of Peirce, considered to be a formation of an "explanatory" hypothesis [9:96], this description does not necessarily fit our inferences. Both peculiar characteristics also apply to the syllogism below.

Syllable complexity in terms of number of phonemes per syllable plays a role in our so far most robust regularities in a sample of meanwhile 51 languages: Languages showing a relatively high syllable complexity use a relatively low number of syllables per sentence [10] and per word. It also plays a role in the following syllogism [6] that makes a step from our metric variables to non-metric categories of classical typology. The inference proceeds from three premises (E, D, B): E and D share the middle term "words per clause", D and B the middle term "syllables per word":

E: if more isolating, then more *words per clause* (is in line with the isolating tendency towards "one morpheme per word")
D: if more *words per clause*, then fewer *syllables per word* (because of a restricted duration of clauses)
B: if fewer *syllables per word*, then more complex syllables (is a paraphrase of B in our first syllogism)
F: if more isolating, then more complex syllables

The formation of new hypotheses and the respective detection of additional, statistically relevant variables increase a theory's explanatory power [11] as well as its predictive success that is the true touchstone in the evolution of knowledge systems [12]. A given body of nomological knowledge about a certain system obviously embodies the potential for such enlargements – through abductive inferences, without recourse to other domains, and sometimes even without any deductions from more general "covering laws".

REFERENCES

1. Schurz, G. (2008). Patterns of abduction. *Synthese 164*, 201-234.
2. Popper, K. (1976 [1935]). *Logik der Forschung.* Tübingen: J.C. B. Mohr.
3. Popper, K. (2002 [1963]). *Conjectures and Refutations: The Growth of Scientific Knowledge.* London/New York: Routledge.
4. Aliseda, A. (2006). *Abductive Reasoning.* Dordrecht: Springer.
5. Reichenbach, H. (2006 [1938]). *Experience and Prediction*. Indiana: University of Notre Dame.

6. Fenk-Oczlon, G. & Fenk A. (1999) Cognition, quantitative linguistics, and systemic typology. *Linguistic Typology 3-2,* 151– 77
7. Von der Gabelentz, G. (1901). *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Tauchnitz.
8. Fenk, A. & Fenk-Oczlon, G. (1993). Menzerath's law and the constant flow of linguistic information. In R. Köhler & B. Rieger (Eds.) *Contributions to Quantitative Linguistics*. Dordrecht: Kluwer Academic Publishers, 11 – 31.
9. Pape, H. (Ed., 1983). *Charles S. Peirce: Phänomen und Logik der Zeichen*. Frankfurt/Main: Suhrkamp.
10. Fenk-Oczlon, G. & Fenk, A. (2010). Measuring basic tempo across languages and some implications for speech rhythm. *Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH 2010)* in Makuhari (Japan). Lisbon: ISCA, 1537-1540.
11. Greeno, J. G. (1970). Evaluation of statistical hypotheses using information transmitted. *Philosophy of Science 37*, 279-293.
12. Fenk, A. (2008). Occam`s razor in the theory of theory assessment. In A. Hieke & H. Leitgeb (Eds.) *Reduction and Elimination in Philosophy and the Sciences. Contributions of the Austrian Ludwig Wittgenstein Society, Vol. XVI*. Kirchberg am Wechsel: The Austrian Ludwig Wittgenstein Society, 89-91.