

## Abstracting in Information Science: Perspectives from Text Linguistics

*Alka Krvavac*  
*Faculty of Philosophy*  
*University of Zagreb*

In the future all world information will be available on-line. The on-line Public Access Catalog (OPAC) is a computerized replacement for the traditional card catalog and information users will want on-line access to all information including a variety of abstracting databases. These bases will ease users' information overload. The main purpose of abstracting is to construct representation of published items in a form suitable for inclusion in some type of database. For this purpose abstracts are texts which should be formally well organized, topically relevant and linguistically adequately verbalized. In this respect, language should be viewed less from linguistic perspective, in the sense of grammar, and more as a complex mental activity verbally realized by a text. The linguistic discipline which operates of these assumptions is text linguistics. Its 7 standards of textuality can contribute to the efficient deciphering of the text stored in a database.

### 1-01

In the future all world information will be available on-line (Mathews 1994). The on-line Public Access Catalog (OPAC) is a computerized replacement for the traditional card catalog and information users will want on-line access to all information including a variety of abstracting databases. These bases will ease users' information overload. The main purpose of abstracting is to construct representation of published items in a form suitable for inclusion in some type of database. For this purpose abstracts are texts which should be formally well organized, topically relevant and linguistically adequately verbalized.

## 2-01

The two scientific disciplines which can contribute with their organized knowledge to the efficiency of retrieval of specific written information are information science and text linguistics. They both have the common subject matter: decoding the information from a linguistically structured written text. They overlap in their field of research which is the intellectual process of judging the content of a document/text. We are referring to the presentation of the conceptual content of the document/text and its verbal realization. Or, linguistically speaking we are talking about transformation of propositions into the concrete textual form which has its topic and linguistic form.

## 3-01

Information science may be interpreted very broadly as the science which deals with storage, retrieval and transmission of information. As Shera (1971) puts it »we have been transferring information all our lives without knowing it like Moliere's 'bourgeois gentilhomme'«. Since language is a major vehicle for the transfer of information, information science may be regarded as chiefly concerned with linguistically encoded written information. It deals primarily with records or documents of one sort or another and the abstract as a text of secondary publication falls within this category. For the purpose of better understanding the purpose and function of this type of text some historical data are given.

## 3-02

It is interesting to note that although one might think that a scientific abstract is rather new in the history of the scientific publications Skolnik (1979) gives some historical data in which he shows that abstract has been already known in Summeric civilisation (3600 B.C.). Skolnik mentions scientists and »scriptors« who were dealing with copying, summarizing and »macrostructuring« books. They were making notes of the most important items for their rulers. Even in later centuries documents of different kinds and functions were copied and summarized.

### 3-03

The first scientific abstract in today's sense of the word dates back to 1662 when Royal society of London was founded. The main task of the society was to back up the development of scientific research and to publish its results. Unfortunately, the idea of regular scientific publications in which abstracts were supposed to summarize the new scientific achievements was realized only 200 years later. In modern times scientific research is published in different periodicals and the abstract as a short informative text falls within the category of periodical's standardization.

### 4-01

Text linguistics, on the other hand, has its own path of development and it is also a relatively new science. Its status is not yet clearly defined as there are different approaches to text analysis. Text linguistics as science started in seventies with de Beaugrande's and Dressler's (1972) »Introduction to Text Linguistics«. In this book the authors stated that text linguistics is trying to solve some problems that structural linguistics could not. Linguistic analysis earlier had not gone beyond the limits of the sentence and the methods known had not allowed to describe relationships between sentences. Supra-sentential regularities could not be reconstructed by traditional morphological and syntactical tools. It was concluded that reliable interpretation should be based on the description of text structure. The problems of supra-sentential relations and of text interpretation in literature had great influence on the further development of text linguistics. Philosophical logic has also exerted a considerable influence especially on the notion of text coherence, inference, anaphoric and cataphoric reference.

### 4-02

Text linguistics operates on the assumptions of seven standards of textuality (de Beaugrande, Dressler 1981).

#### 4-02.1

The first standard is cohesion and it concerns the ways in which the components of the surface text are mutually connected. Cohesive ties are the text elements which

in different ways can tie a textual structure. The most frequent are anaphora for the previously mentioned item in text and cataphora for the item which is going to be mentioned. Then we have lexical elements which are grammatically connected and connectors which show the relationship within a sentence and between sentences.

#### 4-02.2

The second standard is coherence which is concerned with the way the logic of the text underlies the texture. The logical ties can be of various types such as cause, time, purpose and in the case of abstract just purely macrostructures which follow one another.

#### 4-02.3

The third standard of textuality is intentionality which is concerned with the producer's attitude and the goal he is trying to achieve.

#### 4-02.4

The fourth standard is acceptability which refers to the user of information and his attitude towards text types.

#### 4-02.5

Informativity is the fifth standard of textuality which concerns the extent to which occurrences of the presented text are expected / unexpected or known / unknown. This standard overlaps with ISO standard on abstracts. The main function of abstract is to state the new and unknown procedures and results in the field of science.

#### 4-03.6

The sixth standard of textuality can be designated as situationality and is concerned with the level of relevancy of the text to the situation in which it occurs.

4-03.7

The seventh standard of textuality is intertextuality which concerns the factors of knowledge of the text in question in relation to other texts to which it is connected. In the case of the text of abstract it is the correlation between globality and specificity of information given in the abstract in correlation to its article.

5-01

Now let's go back to the beginning. As we have said in the introduction, there is a tendency of all world information to be available on-line. Hence, the need for relevant abstracting databases in the information science sense of the word will be of the outmost importance. The authors, the scientists will sell the results of their work to the international audience and submit the level of their work to the evaluation of international community. On the other hand, the users of abstracting databases will need the integrated specific professional knowledge of the topic in content, information science knowledge in the sense of the abstracts' standards, functions and use and linguistic knowledge in the sense of mental activity in which words participate. For this purpose the integrated knowledge of the disciplines of information science and text linguistics which have only been outlined here can contribute to the level of predictability precision of the information in the article and to the cost-effectiveness of retrieval, which is a must of today's development in technology.

REFERENCES:

- BEAUGRANDE, Robert de & WOLFGANG Dressler (1981). *Introduction to text linguistics*. Longman.  
Documentation-Abstracts for publications and documentation (1976). ISO. 214.  
HENRY, Charles (1993). »The surface of language and humanities computing«. *Computers and the humanities*, 27, 315–322.  
HINES, Andy (1994). »Jobs and infotech«. *The futurist*, Jan. -Feb.  
JONES, Karen Spark & MARTIN Kay (1973). *Linguistics and information science*. Academic press.

- KRVAVAC, Alka (1988). »Procedure in abstract writing«. *Termlet News, Journal of the international network for terminology*, 19, Vienna: Infoterm, 1–16.
- KRVAVAC, Alka (1990). »Analiza jednog tipa apstrakta kao znanstvenog teksta«. *Projekt pragmatolingvistika prirodnih jezika*. Zagreb: Filozofski fakultet. 1–10.
- KRVAVAC, Alka (1991). *Utjecaj prezentacije sažetka na razumijevanje teksta*. (neobjavljena doktorska dizertacija). Zagreb: Sveučilište u Zagrebu.
- KRVAVAC, Alka (1993). »Testiranje razumijevanja apstrakta kao teksta u znanosti«. *Strani jezici*, XII, 2, 93–101.
- KRVAVAC, Alka (1994). »Abstract as a text in science with respect to its information relevance« In: *Applications and implications of current LSP research*. LSP Bergen, Norway, 93, Vol. I. 746–751.
- KRVAVAC, Alka (1995). »Specialist communication via database abstract: A Croatian and English case study«. *The 10<sup>th</sup> European Symposium. on Language for Special Purposes*. Vienna, Austria.
- LANCASTER, F. W (1991). *Indexing and abstracting in theory and practice*. Graduate school of library and information science. University of Illinois. USA.
- MATHEWS, Joseph R (1994). »The distribution of information: the role of on-line public access catalogs«. *Information services and use*, 14, 73–78.
- RIESSER, Hannes (1978). *On the development of text grammar*. In: *Current trends in textlinguistics*. Ed. by Wolfgang V. Dressler. Berlin–NewYork: Walter de Gruyter.
- SHERA, Jesse H. (1971). »The sociological relationships of information science«. *Journal of the ASIS*. 22. 77.

## APSTRAHIRANJE U INFORMACIJSKOJ ZNANOSTI SA STAJALIŠTA LINGVISTIKE TEKSTA

U budućnosti će sva svjetska informacija biti pohranjena na *on-lineu*. *On-line* javni katalog je kompjutorski nadomjestak za tradicionalni bibliotечni papirni katalog i korisnici informacija moći će imati pristup svakoj *on-line* informaciji, pa tako i bazama apstrakata. Te će baze olakšati informatički *overload*. Glavni cilj apstrahiranja je konstrukcija prezentacije publicirane jedinice u formi koja odgovara organizaciji baze podataka takvog tipa. Za tu namjenu apstrakt je tekst koji formalno mora biti dobro organiziran, tematski relevantan i lingvistički adekvatno verbaliziran. U tom kontekstu jezik se mora promatrati manje s lingvističke perspektive, u gramatičkom smislu, a više kao kompleksna mentalna aktivnost koja je verbalno realizirana tekстом. Lingvistička disciplina koja djeluje na tim pretpostavkama je lingvistika teksta. 7 standarda tekstualnosti mogu pridonijeti uspješnom dešifriranju teksta pohranjenog u bazi podataka.