

Information Processing of Text in the System of Teaching Languages at Technical HEIs

Ospanova Bikesh Revovna¹, Azimbayeva Zhanat Amantayevna¹ & Timokhina Tatyana Vladimirovna¹

¹ Karaganda State Technical University, Mira Boulevard, 56, Karaganda, Kazakhstan

Correspondence: Ospanova Bikesh Revovna, Karaganda State Technical University, 100027 Mira Boulevard, 56, Karaganda, Kazakhstan.

Received: July 9, 2014 Accepted: September 15, 2014 Online Published: October 30, 2014

doi:10.5539/ass.v10n21p275

URL: <http://dx.doi.org/10.5539/ass.v10n21p275>

Abstract

The present day level of the society informatization makes high demands on the level of information competence of specialists in all spheres, and one of the basic abilities lying in the basis of any professional activity is the ability to work with the text material. Today a specialist has to process great volumes of professional information. When working with special literature: scientific, technical, economic, juridical and other, a specialist is not only to quickly and correctly orient himself in the content but to possess the art of the information text processing. The aim of the text information processing is extracting useful and valuable information concerning a concrete problem, interpreting the content into more or less detailed form, depending on the practical value of information and its further use needs.

One of the key competences that are formed in practical classes at language departments when working with texts is a communicative competence supposing student's ability to interact with a text, to perceive information and to interpret it adequately, to develop an own statement. The problems arising in forming this competence are connected with the gap between the acquiring of theoretical base and forming of practical abilities and skills. Thus, the novelty of the study consists in the fact that there was undertaken an attempt to develop a methodology of working with a text in practical classes; it is necessary to teach working with information in the texts: to form a competence that supposes the ability to extract information, to analyze, to transform it, to use for solving problems.

Keywords: information, text, linguistic study, processing, competence, specialist, language, functioning, laws, linguistics

1. Introduction

In the Conception of reforming and modernizing the Republic of Kazakhstan education system for the period till 2020, there is underlined the necessity of developing a model of the future education permitting to form a mobile, dynamic, constructive specialist who knows his language and culture, the world languages, and is able to make independent decisions in the situation of choice. In this connection there arises the matter of improving specialists training quality, the importance of the education content and technology refreshment, particularly the language one.

Since students of technical HEIs possess insufficient abilities permitting to satisfy their information needs, the search for new technologies in language training of technical HEIs graduates attracts attention of native and foreign scientists.

In present day conditions of continuous education, at the level of the latest achievements in the field of linguistic science, there takes place the reconsideration of the information analysis procedure. The development of information technologies opens huge prospects for specialists in the field of linguistic studies. A text is the basic object of a linguistic study. "The main, general aim of developing a text (any one) is information. Any text contains some information. The general amount of information contained in the text is presented as its information intenseness" (Valgina, 2003).

From the linguistics point of view (I.R. Galperin's theory), the understanding of a text is the reading of various types of text information: "substantial-factual" (information of facts, events...), "substantial-subtextual" (latent information), "substantial-conceptual" (author's thoughts, emotions that are reflected in the text reckoning upon

readers' understanding) (Galperin, 2007). In the works of such scientists as Granik, G.G., Shapovalov, S.A., Sobolev, O.V. (Granik, et al.), there is defined the problem of understanding text and its solution ways. In publications of the PISA-2009 international study, in the site "Center of education quality assessment IOSO RAS", there are opened modern approaches to overall understanding, the connection of understanding with education, creative work (Publication of the international study PISA, 2009).

The problem of our study consists in that students of technical HEIs cannot recognize specific information; they do not take into consideration its adequacy in relation to the defined objective. The process of perceiving the text information is characterized by fragmentariness and incompleteness.

But in order to live adequately in the information society, a person is to process a huge amount of information. The avalanche of information contained in information flows cannot be analyzed by a person and processed by him, and of great help here can serve the latest technologies where there takes place the automatic processing of the text material. Only new and useful information possesses some value, and it is an indicator of the text informativeness because the text information intenseness is always the absolute indicator of the text quality.

2. Main part

In the course of study there were used the following methods: general-scientific methods of system analysis, method of quantitative analysis, information analysis, comparative method for revealing text characteristics.

3. Results of the study

The work with a text is the main method of obtaining information, therefore the ability to perceive and to understand the text content defines the progress in studies at technical HEI. The condition of achieving the present day level of humanitarian-technical education is a dialogue with a text that is adequately perceived by a student.

Based on the carried out study of the text information processing there were revealed the basic abilities of the text perceiving connected with the reading of information (factual, subtextual, and conceptual):

1. understanding the general content of the text, the main sense;
2. ability to actively perceive text information;
3. ability to extract factual information, i.e. to separate the main information from the secondary information (main from additional);
4. ability to fill in semantic gaps, to reveal latent information and to make conclusions based on apparent and latent information;
5. ability to extract conceptual information, to interpret and generalize it;
6. ability to compare information;
7. ability to find in the text a confirmation of the statement, to differentiate a subject ("opinion") from a fact, to prove one's statement with an example from the text;
8. finding key words, word combinations, sentences (important informatively), realizing the thematic and semantic unity of the text;
9. ability to orient in the text structure, to trace the logics of a statement;
10. ability to use various methods of reading: viewing, searching, oriented to the selection of information needed, analytical);
11. ability to keep information in memory (to compare, to combine fragments into a general information picture);
12. ability to reformulate information, to compress the text;
13. ability to state one's personal opinion (to reflex on the text content and form).

4. Discussion of the Study

The text information processing is one of the most widely spread written forms of information extraction and its semantic processing at the advanced or even final stage of language teaching. The basis of such activities is the complete understanding of the text, i.e. the ability to work with scientific-technical literature.

Alongside with good knowledge of a language for successful work at text information processing a specialist needs reading scientific literature on the problems of his field of activity, specialty and to orient freely in publications, to be aware of the present day state of the problems considered, to know the country technical

achievements. A student is to be able to separate the main things from the secondary ones, to analyze and generalize the material processed.

The essence of the text information processing consists in a brief and generalized stating the original source content (the primary text) in accordance with the tasks. The aim of the text information processing is the extracting of useful and valuable information on a concrete problem presenting a statement interest.

The terms “primary” and “secondary” texts appeared in applied linguistics from the theory of scientific-technical information, i.e. informatics. The conceptions “primary” and “secondary” are used as a basis of information documents classification. The secondary texts are made as a result of transforming an original, primary, basic text.

A necessary condition for working with a text is a preliminary understanding of the primary text on the whole (there is required gradual understanding of individual fragments of the text). Only under such a condition a student can find in it the main, essential information and to omit the secondary information. A secondary text is to be considered as a new one, which forming is determined by the result of overall text understanding.

Extraction of information from the text is connected with a new approach to the text processing in the aspect of communication, namely, the text grammar arising that is called the text syntax or the text linguistics. Methodological intellectual work with the text material supposes the acquaintance with the structure of the entire text and its components. The knowledge of the text linguistics elements gives a possibility to come to the actions of the text shortening and revealing the main content.

The next important stage is attentive, repeated reading of the text with separating the key fragments. A process of key fragments finding permits to watch the developing of the main themes of the text in the structure of a paragraph and groups of paragraphs. All the further work at the text information processing is carried out based on the chains of fragments that make the semantic frame of the text.

An important method presenting a difficulty in the text information processing is the work at paragraphs with the implicit sense, i.e. supposed and removed from the general content. The implicit subject is most often reduced to that in the primary text there are mentioned some facts that have a private character and cannot be called the key ones, however, they have an important semantic load. In such cases it is necessary to find a generalizing word, word combination or a sentence that would incorporate the sense of these formulations and synthesize their general sense.

Separation of the key fragments makes the first stage of the text information processing. The next stage is organization and regrouping of the separated fragments in accordance with their thematic orientation. Then there follows creation of the text logical plan. The separated key fragments are written down (or underlined) in the order, in which they were met in the primary text. Then they are grouped according to the thematic principle around some large subsubjects developing the main theme of the text.

The problem of the text understanding is one of the urgent problems of students’ educational and scientific activities at higher education institutions. Any text possesses an internal and external structure. The text that is a product of the speech activity, the basic communicative unit of speech can be considered as one of the effective means of forming skills and abilities of scientific communication. Being a communicative-cognitive unit of the professional communication mechanism, the text permits to communicate with specialists. It is well known that every text is characterized by the unity of the subject and plot, its relative completeness, coherence, integrity, internal structure, pragmatic directive.

In the linguists’ common opinion, the basic function of the text as a product of the communication act is sending and receiving the information that presents a “result of reflecting real phenomena and objects, complex relations of phenomena and situation”. At present, the problems of information are being developed rather intensely in various fields of knowledge: linguistics, mathematics, philosophy, logics, and psychology. As a result, this concept is widely interpreted. In linguistics, a lot of researchers separate two basic types of information: objective, or objective-logical relating to statements of the reality objects, and subjective, or expressive-emotional-estimating, presenting the expression of the speaker’s attitude to the reality facts and relating to their subjective estimation (Slyussareva & Teplitskaya, 1978).

It is known that the concept of “the text subject” is interpreted in a wide and narrow aspect. A wide aspect means the information contained in the text; a narrow aspect means the designation of the object spoken about. The narrow aspect of the subject interpretation is adopted in the theory of the statement actual segmentation. We agree with such interpretation and keep this understanding of the subject. The interest to the text studying is caused by the will to explain a language as a global phenomenon from the point of view of modern linguistics, as

an integral means of activities realized throughout the text. The interest to the text is also explained by the scientists' aspiration to explain the laws of the language functioning in the aspect that is larger than a sentence.

The search of approaches to the problem studying is possible owing to theoretical preconditions developed in the studies of home and foreign scientists, such as Galperin, I. R. (1981), Zvegintsev, V. A. (1980), Gindin, S. I. (1971), David J. C. & MacKay (2003), Willie J. R. Martin (Willey et al., 1983), Frank Smadja (Smadja, 1993), Gross E.U. (Grosse, 1976), Tomas C. Landauer (Thomas, Landauer, & Harshman, 1990), David Klahr (Klahr, 1989), Saket Mengle and Nazli Goharian (Mengle & Goharian, 2009), William P. Eveland, Jr., Sharon Dunwoody (Eveland, et al.), Hall, R.H. (Hall, 2009). In the works of foreign and home researchers there are considered a lot of conceptions as for the information model of the text.

Let's consider in brief the most fruitful propositions of present day studies of the text perceiving and understanding. The starting point is the fact that a "coherent text is more than a language itself and much more than a sequence of individual sentences" (Graesser, 1997).

A lot of researchers believe that any text means some content on a certain subject; texts differ by the amount of information. In the texts of a large volume the subject unity is demonstrated by the subject levels. In such texts the leading subject is divided into a number of the component subsubjects that are segmented into microsubjects, which are considered minimum units of the speech sense. Any text possesses semantic and structural completeness, composition form, various types of lexical, grammar, logical and stylistic relations between the text elements, information intenseness, different methods of the idea presentation, etc. From this perspective, the text can act as an independent object of study. "As a rule, this is accomplished on the basis of partial text parsing, which allows performing the processing of news flow coming from news agencies. Since this problem is quite complex, both from theoretical and technological point of view, creation of significant systems of information extraction is feasible within the commercial companies" (Caldeira et al., 2005).

The sphere of professional activity is served by a special language, the language of professional communication. The structure of this language is irregular, there are separated several language seams depending on the extent of codification and information intenseness: terminology, oral professional vocabulary (professionalisms), professional slang, nomenclature. The national science considers communication traditionally as information exchange, while the west researchers, such as M. Braunstein, for example, emphasize the behavioral and axiological aspects of communication (Marty, 2001). With time, the definition of communication as a purely information process exhausted itself, since scientists paid attention to the loss of a part of information in the course of communication.

Teaching students the skills of scientific texts understanding and writing scientific works on a language specialty is the main problem in the training of a highly-educated specialist possessing the language bases and its practical use. Technical HEIs graduates are to include in their system of the functioning knowledge not only technological but also linguistic literacy. Specialist's professionalism will be revealed when he can freely operate with the terms used in his specialty. For this it is necessary to form at technical HEIs graduates the "language sureness" in all kinds of the speech activities using effective speech models.

In the number of the basic problems of language teaching at technical HEIs there is professional communication and the quickest assuring of the complete acquiring of the specialty language base. The conception of teaching the scientific text understanding based on the modeling of semantic relations was developed for simplifying the text information development. Modeling means the substitution of studying a real concept of interest by studying a similar phenomenon on a model. The text modeling is needed only there, where the problem of misunderstanding exists.

Thus, the development of the text linguistics and changing the approaches to the linguistic education define in a new way the role and place of the text in teaching languages. Today the text is one of the most important components of the training content, particularly of teaching the language of specialty. The latest achievements of linguistics and methodology of language teaching permitted to deepen the understanding of the role and place of the scientific text when working at the language of specialty. The language of specialty is offered to students in the form of separate training, logically ordered texts. A training scientific text is used as a unit in which there is convincingly revealed the studied object content, there is given the totality of its features in a certain aspect of consideration.

5. Conclusion

Present day conditions of studying at a technical HEI force to verify the structure of the aspects of profession-oriented texts' information processing. There is no doubt that the text material in specialty is not to

repeat the courses of special disciplines in the entire volume. It is necessary to define the main bodies of professionally relevant texts for each discipline that reflect characteristic features and interaction of various types of information.

It is necessary to develop a typology of the texts system taking into account student's speech activity in classes in specialty which optimizes the teaching process and stimulates the interest to language learning. Such an approach permits to extract and process the relevant information, which finally makes the basis of information activities in the field of technical sciences and is a necessary condition for specialists' successful work.

It should be noted that a text manifests new qualities of the learnt languages and is able to represent the highest level of the students' cognition opening the way to perfecting speech abilities and skills. As a didactic unit the text will permit to merge the two most important tendencies in language learning, both native and foreign: cognition of the language system and cognition of norms and rules of communication in various situations of life.

Thus, teaching the extraction and understanding of information supposes students enriching with special knowledge and problems of specialty, awareness of its benefit and necessity for the further professional communication. This all finally permits to form the language base of students' training.

Specialists' training at technical HEIs consists in forming communicative abilities that would permit to perform professional contacts in various spheres and situations, forming the future specialist's aspiration and ability to function as a strong linguistic personality of democratic type, possessing a high linguistic competence in the field of Russian, Kazakh and English languages, in professionally relevant speech events of various types, in different forms, styles, genres of profession-oriented speaking-and-thinking activity.

6. Summary

The aim of this article is an attempt to develop a methodology of applied character for working with the text material for the information text processing. Within the frames of the study carried out we hope that this will serve a needed pulse for other authors for the purpose of their scientific search in this interesting and practically significant field of knowledge.

Thus, the situation at the internal and international labor market requires a high level of knowing the language of professional communication, which in turn is an obligatory condition for specialists' professional work and scientific growth. So, the analysis of the text information processing permits to form optimal conditions for specialist's information culture, his thesaurus in the system of higher technical education.

References

- Caldeira S. M. G., Petit Lobao, T. C., Andrade, R. F. S., Neme, A., & Miranda, J. G. V. (2005). *The network of concepts in written texts*.
- Eveland, Jr. W. P., & Dunwoody, S. (n. d.). Examining Information Processing on the World Wide Web Using Think Aloud Protocols. *Mediapsychology*, 219-244.
- Galperin, I. R. (1981). *Text as an object of linguistic study* (p. 25).
- Galperin, I. R. (2007). *Text as an object of linguistic study*. M.: Komkniga.
- Gindin, S. I., (1971). Two principles of internal organization of texts and essence of the conception "text cohesion". *Semiotic problems of languages of science, terminology and informatics*.
- Graesser, A. C. (1997). *Discourse comprehension*. *Ann. Rev. Psychol.*, 48.
- Granik, G. G., Shapoval, S. A., Kontsevaya, L. A., & Bondarenko, S. M. (n. d.). *Literature. Let's study to understand belles-lettres text. Practical problems*. Retrieved from <http://www.alleng.ru/d/lit/lit13.htm>
- Grosse, E. U. (1976). *Text and Kommunikation*.
- Hall, R. H. (2009). *Information Processing Theory*. [Electronic Version]. Missouri University of Science & Technology.
- Klahr, D. (1989). Information-processing approaches. *Annals of child development*, 6, 133-185.
- MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.
- Marty, B. (2001). *Communicating Effectively for Dummies* (p. 356). N. Y.: Wiley Publishing, Inc.
- Mengle, S., & Goharian, N. (2009). Passage Detection using Text Classification. *Journal of the American Society for Information Science and Technology*, 60(4), 814-825. <http://dx.doi.org/10.1002/asi.21025>

Publication of the international study PISA-2009. Retrieved from <http://www.centeroko.ru/>

Slyussareva, N. A., & Teplitskaya, N. I. (1978). *Hyper-syntactic level of language and linguistic fragmentation of texts* (pp. 96-97). Sentence and text in semantic aspect. Kalinin.

Smadja, F. (1993). Retrieving collocations from text: Xtract. *Computational Linguistics*, 19(1), 143-177.

Thomas, K., Landauer, T., & Harshman, R. (1990). Latent semantic analysis. *J. Amer. Soc. of Information Science*, 41(6).

Valgina, N. S. (2003). *Theory of text* (p. 279). M.: Logos.

Willey, J. R., Martin, B. P. F. Al, & van Sterkenburg, P. J. G. (1983). On the processing of text corpus: From textual data to lexicographical information. In R. K. Hartmann (Ed.), *Lexicography: Principles and Practice* (pp. 77-87). Academic Press Inc., London, U.K.

Zvegintsev, V. A. (1980). *Text units' all-statefulness* (pp. 13-21). Bull. of USSR AS. Ser. Lit. & lang. No 1.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).