APPLIED LINGUISTICS, COMPUTATIONAL LINGUISTICS AND OTHER INTERDISCIPLINARY SCIENCES AS EFFECTIVE INTERCULTURAL COMMUNICATION STRATEGIES

Serebrianska Irina,
Doctor of Philological Sciences, Associate Professor
ORCID ID 0000-0002-1720-4873
Nizhyn Mykola Gogol State University, Nizhyn, Ukraine
serebryanska@ukr.net

The article deals with the issues of interaction of linguistics and modern computer information technologies. It systematizes the actual problems and ways of their integration based on the analysis of the latest scientific works and the appropriate educational programs in higher education institutions all over the world. It draws attention to interdisciplinary sciences such as applied linguistics, computational linguistics, linguistic informatics, quantitative linguistics, mathematical linguistics, Internet linguistics, and integrated knowledge. The cooperation of two different areas is very productive and prospective. It contributes to the development of artificial intelligence and provides wider professional opportunities, effective intercultural communication and further improvement of society. The representatives of two, very different fields become mutually beneficial and complete each other successfully. The result of this interaction is a universal specialist which is in great demand nowadays. The main issues of the interaction of linguistics and modern information technologies are the following: 1) the role of virtual space and modern information technology in the development of linguistics; 2) the role of linguistics in the development of virtual space and information technologies, and the training of computer specialists; 3) the place of virtual space and modern information technologies in preparing linguists and translators (machine translation, translation systems, etc.); 4) the emergence of interdisciplinary sciences and educational courses at the intersection of two areas to provide society with professionals with integrated knowledge; duplication of their conceptual and terminological apparatus and research methods (due to their novelty); methodological support of the integrated educational process.

__Keywords__: interdisciplinary sciences, applied linguistics, computational linguistics, linguistic informatics, quantitative linguistics, mathematical linguistics, Internet linguistics, advanced computer information technologies, intercultural communication.

Introduction

The people of the XXI century and society, in general, are so accustomed to the interactive information environment and the Internet space in both everyday and professional life that they are unlikely to be able to do without them. In such conditions, the relationship of language as a means of communication and linguistics with advanced computer information technologies plays a very important role. It determines the **relevance** of the study.

The **purpose** of the paper is to systematize the actual problems of the interaction of linguistics and modern information technologies based on the analysis of the latest scientific works and the appropriate programs in higher education institutions. In the process of research, general scientific **methods** of analysis and synthesis were used, as well as the method of classification, they allowed to analyze scientific and professional texts and draw conclusions about the subject of the research.

_____

New trends in social development require knowledge in various fields and the appearance of new methods. The use of rapidly evolving computer information technologies allows for achieving better results in any field. For example, for linguists, it is primarily the use of computers in their research: automatic search of bibliographic information, computer processing of linguistic data, electronic corpora of texts and translation, machine and automated translation systems, computer technologies in lexicography. In turn, for professionals in computer information technologies, a high level of language is in demand. Thus, the representatives of two, very different fields become mutually beneficial and successfully complement each other. At present, specialists with interdisciplinary knowledge and skills both in linguistics and in computer science are needed in particular.

The idea of the mutual benefit of computer information technologies and linguistics is complicated and comparatively new. Among early publications in the field is "Readings in natural language processing", dated 1986 (Grosz et al., 1986). The most comprehensive textbook on computational linguistics in the USA is "Natural language understanding" (Allen, 1995). Statistical approaches to natural language processing are described in the edition "Computational Linguistics" (Computational Linguistics, 1993). The collection of papers "Artificial intelligence" (Fernando, Grosz) has a special volume on natural language processing, dedicated to the knowledge-based approach to natural language processing. The other authors of the textbooks on computational linguistics, popular all over the world are Christopher Manning and Hinrich Schütze, "Foundations of Statistical Natural Language Processing" (Manning, Schütze, 1999) and Daniel Jurafsky and James Martin, "An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Second Edition. Prentice Hall" (Jurafsky, Martin, 2008).

Interdisciplinary sciences

As for tandem of advanced computer information technologies and linguistics, some new sciences have been actively formed at their intersection: linguistic informatics, corpus linguistics, computational linguistics, mathematical linguistics, quantitative linguistics, Internet linguistics, etc. All of them convincingly testify to the need to integrate knowledge – on the one hand, technical, computational, mathematical and linguistic (for successful processing of linguistic information); on the other hand – linguistic knowledge and skills combined with technical and digital ones (to improve speech structures of computer specialists, their understanding of basic concepts in the field of computer technologies, to develop and maintain systems with linguistic components).

Nowadays interdisciplinary links are implemented in computer translation of professional texts, in linguistic research and teaching foreign languages, in judicial linguistic expertise, automatic processing of texts in different languages. Technicians also use linguistic methods in their work, for example, studying methods and algorithms of automated formation of logical-linguistic models of textual information; language tools for designing the websites' content, etc. Modern linguists are interested in the following technological issues: informatization of language education, didactic properties and functions of information and communication technologies, information security of students in the process of Internet projects, search and selection of materials from the Internet for learning, blog technology, wiki technology, service Twitter and the linguistic corpus in language teaching (Association).

As for the latest studies, the impact of informatics and the new technology on linguistics is represented in the paper of Roberto Busa "Informatics and new philology" (Busa, 1990). Another article stresses: "In the present day world linguistics became a foundation for developing information technologies. While informatics is a vigorously developing field of human activity, the union of mathematics and linguistics will continue playing its role in the development of science" (Ospanova, 2014). We support the idea that the interaction of linguistics with other sciences (mathematics, biology, cybernetics, informatics, etc.) can be performed only when using the mathematical apparatus.

Some scientific works represent the application of linguistics methods in advanced information technologies, the use of models of information interaction in computational

linguistics, the transformation of theoretical linguistics in the information area (Klimova, 2009; Ghasemi et al., 2011; Chekharin, 2016). They help to form a terminological system for the representatives of the information sphere (Sinyagovska, Kryzhovska, 2020). Lately, the following textbooks have appeared for the training of specialists in the interdisciplinary fields: "Modern information technologies in linguistics" (Petrasova, 2020), "Quantitative linguistics and new information technologies" (Moiseeva, 2017), "Fundamentals of computer science and applied linguistics with a module Copywriting" (Prokopenko, Gartsunova, 2020), "Nanotechnologies and linguistics: Forecasting and perspectives of interaction" (Potapova, 2007) and others. All of them are devoted to the issues of advanced computer information technologies in linguistics and language education, the dynamics of intellectual learning systems, computerization of learning in the humanities, including foreign languages, distance learning concepts, hypertext in linguistics, automated learning systems, the specifics of Internet communication. For example, the textbook "Information Technology and Linguistics of the XXI century" (Guslyakova, 2016), along with practical tasks and information about the specifics of the profession of a linguist, provides information on the main programs of the machine and automated translation and a brief description of popular search engines.

Due to the analyses of the scientific works on the topic we propose the review of the most popular interdisciplinary sciences:

1. **Linguistic informatics** is the study of the principles of building models that reflects the specific processes of speech mental activity in human situations. That is, these are information systems that can be used to model and analyze texts, store and process data and make decisions (Educational and methodical).

2. **Computer linguistics** is an interdisciplinary field concerned with the computational modelling of natural language, as well as the study of appropriate computational approaches to linguistic issues (English Wikipedia). In general, it draws upon several areas – linguistics, computer science, artificial intelligence, math, logic, philosophy, cognitive science, cognitive psychology, psycholinguistics, anthropology and neuroscience. According to "Stanford Encyclopedia of Philosophy" (Computational Linguistics, 2014) computational linguistics is the scientific and engineering discipline concerned with understanding written and spoken language from a computational perspective, and building artefacts that usefully process and produce language, either in bulk or in a dialogue setting. The theoretically justified statement is that since language is our most natural and most versatile means of communication, linguistically competent computers would greatly facilitate our interaction with machines and software of all sorts, and put at our fingertips, in ways that truly meet our needs, the vast textual and other resources of the Internet.

Jerry Hobbs, analyzing perspectives in Computational Linguistics (Hobbs), admits that computational linguists study natural languages, such as English and Japanese, rather than computer languages, such as Fortran, Snobol, C++, or Java.

According to the Linguistic Society of America, there are two aims of computational linguistics: 1) the technological one – to enable computers to be used as aids in analyzing and processing natural language; 2) the psychological one – to understand, by analogy with computers, more about how people process natural language.

We should admit, a scientific study of language from a computational perspective is actual and very productive as computational linguists are interested in providing computational models of various kinds of linguistic phenomena. "These models may be "knowledge-based" ("hand-crafted") or "data-driven" ("statistical" or "empirical"). Work in computational linguistics is in some cases motivated from a scientific perspective in that one is trying to provide a computational explanation for a particular linguistic or psycholinguistic phenomenon and in other cases, the motivation may be more purely technological in that one wants to provide a working component of a speech or natural language system. Indeed, the work of computational linguists is incorporated into many working systems today, including speech recognition systems, text-to-speech synthesizers, automated voice response systems, web search engines, text editors, language instruction materials, to name just a few" (Association).

3. **Mathematical linguistics** is the mathematical discipline which objective is the development and study of ideas forming the basis of a formal apparatus for the description of the structure of natural languages (that is, the metalanguage of linguistics). This direction of science appeared in the 1950s, it is the oldest one among others. According to "Encyclopedia of Mathematics" (Encyclopedia), "it was brought to life first of all by the internal needs of theoretical linguistics, which at that time was ripe for an elaboration of its basic ideas, and also by problems in the automatic processing of linguistic information".

Mathematical linguistics is constantly evolving along the path of transforming into a purely mathematical discipline, being essentially a branch of mathematical logic. At the same time, the circle of applications of mathematical linguistics is expanding; and its methods have found application in the theory of programming. Mathematics penetration into linguistics is especially deep in connection with searching for a solution to linguistic analysis problems at the level of information systems. The scientists truthfully admit that using mathematical methods in different scientific-practical fields such as linguistic synergy, the theory of vocal communication, linguistic didactics, the theory of script, theory of translation allows to increase the accuracy of the results and give a possibility of a more objective explanation of language phenomena, to modernize the system of higher vocational education of mathematical and humanitarian trends, define the prospects of scientific research at the intersection of the fields of hard sciences and humanities (Ospanova, 2014).

4. **Quantitative linguistics** is a sub-discipline of general linguistics and mathematical linguistics. Quantitative linguistics deals with language learning, language change, and application as well as the structure of natural languages. It investigates languages using statistical methods. Its most demanding objective is the formulation of language laws and, ultimately, of a general theory of language in the sense of a set of interrelated languages laws. Synergetic linguistics was from its very beginning specifically designed for this purpose. Quantitative linguistics is empirically based on the results of language statistics, a field that can be interpreted as statistics of languages or as statistics of any linguistic object. This field is not necessarily connected to substantial theoretical ambitions (English Wikipedia).

5. **Internet linguistics** studies new language styles and forms that have arisen under the influence of the Internet and other new media, such as Short Message Service (SMS) text messaging. Since the beginning of human-computer interaction leading to computer-mediated communication and Internet-mediated communication, experts, such as Gretchen McCulloch have acknowledged that linguistics has a contributing role in it, in terms of the web interface and usability. Studying the emerging language on the Internet can help improve conceptual organization, translation and web usability. Such a study aims to benefit both linguists and web users combined (English Wikipedia).

English linguist D. Crystal in his book "Language revolution" (Crystal, 2001) rightly called the Internet occurring the most significant fact that led to revolutionary transformations in the language. The main perspectives of Internet linguistics are sociolinguistics, education, stylistics and applied linguistics.

However, there are terminological problems connected with new sciences. Some scientists, unfortunately, sometimes confuse the concepts of "linguistic informatics" and "applied linguistics", including the first science into the second one (Franchuk, 2019). We believe that it is better to distinguish them based on different keywords: "informatics" in the first case and "linguistics" in the second one. We consider linguistic informatics a science about language transformations, in particular, the modelling of transitions from the complex multi-component and layered system of natural language to the language of binary codes (analysis); and the opposite direction – from languages of machine level – to high-level languages and natural language (synthesis) (Kozerenko, 2006). Another example is the use of nominations "computational linguistics" and "linguistic informatics" (Yatsko, 2014). Confusion of these concepts is caused by their novelty and it needs to be solved.

Interdisciplinary courses and educational programs

The role of computers in language teaching has changed significantly in the last 30 years. It allows "to more fully integrate computer technology into the language learning

process. Multimedia programs incorporating speech-recognition software can immerse students into rich environments for language practice. Concordance software with large language corpora provide students with the means to investigate language use in authentic contexts. And the Internet allows for a myriad of opportunities to communicate in the target language, access textual and multimedia information, and publish for a global audience" (Warschauer, Healey, 1998).

The emergence of the following interdisciplinary courses in the educational process of higher education institutions are in trend: "Computer word processing", "Computer lexicography", "Computer linguodidactics", "Linguistic informatics", "Information technologies in linguistics". They are very important for an innovative society. The appropriate educational programs are represented on the sites of the world's universities ("Applied linguistics", "Computational linguistics", etc.).

The future requires qualified specialists with universal competencies, computational linguists, who can develop algorithms for recognizing oral language, QA-systems, machine translation systems, text generators, electronic dictionaries and databases, information retrieval systems, and work with SQL, language processing technologies, and programming. Such activities aim at artificial intelligence, machine learning, Big Data and, of course, they are socially oriented. Our analysis allowed us to determined the most spread interdisciplinary programs in the world, described below. One of the highly-rated educational institutions – Trinity College (Dublin) proposes an integrated, interdisciplinary program "Computer Science, Linguistics and a Language", which is on offer (Trinity College). Students learn computer science, study linguistics, the scientific study of language and speech, and study a specific language (with a choice of French, Spanish or Irish). It emphasizes the following questions: how these disciplines intersect; computational, empirical approaches to the study of language; developing language proficiency and acquiring the knowledge and skills that are important in the growing field of speech and language technologies (machine translation, speech synthesis and recognition). The degree combines subjects on modern information technologies (a comprehensive introduction to computers; how to program them and certain fundamentals of how they work; more advanced programming; a study of fundamental data structures and algorithms and specifically an introduction to computational treatments of language: computational linguistics) with linguistics modules including human communication. These encompass both theoretical and computational elements (such as Speech Science and Phonetics, Computational Morphology).

The University of Groningen has developed the program "Language and Communication Technologies" as a part of the international Erasmus Mundus program (University of Groningen). It combines Theoretical Linguistics and Computer Science. Another example is in The Uppsala university where the Master's Program in "Language Technology" teaches how to make computers perform useful and interesting tasks involving human language (Uppsala Universitet). It focuses on both humanistic insights into the way language works and on the use of computational methods.

The University of Washington launched the Master of Science in Computational Linguistics program in 2005 (University of Washington). It is oriented for the tech industry in the area of natural language processing. It prepares new professions such as language engineers, computational linguists and artificial intelligence engineers.

Conclusions

The interaction between linguistics and advanced computer information technologies is represented in theory and practice. The main issues are the following: 1) the role of virtual space and modern information technology in the development of linguistics (hypertext technologies of text representation, lexicography, terminology and terminography, text as a system, linguosynergetics, etc.); 2) the role of linguistics in the development of virtual space and information technologies, and in the training of computer specialists; 3) the place of virtual space and modern information technologies in preparing linguists and translators (machine translation, translation systems, etc.); 4) the emergence of interdisciplinary sciences and educational courses at the intersection of two areas to provide society with professionals

with integrated knowledge; duplication of their conceptual and terminological apparatus and research methods (due to their novelty); methodological support of the integrated educational process.

The development of these integrated fields contributes to the development of artificial intelligence. From the technological perspective, there are three uses for natural language in computer applications: 1) natural language interfaces to software; 2) document retrieval and information extraction from a written text; 3) machine translation. The significant problem in processing natural language is ambiguity. Among others, there is a need to determine the terminology of new scientific areas. Future developments in networked communication, multimedia, and artificial intelligence will likely converge, creating a potentially more central role for interdisciplinarity. Modern education provides the appropriate needs of society with new knowledge and skills.

The results of the interaction of two analyzed areas are increasingly discussed at the global level: at international conferences on artificial intelligence (i.e. "Linguistic Informatics – State of the Art and the Future" since 2005) and on computational linguistics (i.e. "KOLING"). The journal "Computer Linguistics" is published in the USA. The Association for Computational Linguistics is also functioning and developing the corresponding directions. In general, unity of linguistics with advanced computer information technologies provides broader professional opportunities for linguists and translators. It becomes the background for stable and productive intercultural communication and the further development of our society.

REFERENCES

Grosz at al., 1986 – *Grosz, Barbara J., Karen Sparck Jones, and Bonnie Lynn Webber* (1986). Readings in natural language processing, Santa Monica, CA: Morgan Kaufmann.

Allen, 1995 – *Allen, James F.* (1995). Natural language understanding. Benjamin Cummings Pub. Co. 2nd edn.

Computational Linguistics, 1993 – *Computational Linguistics* (1993), vol. 19, no. 1, March: Special issue on using large corpora: I.

Fernando, Grosz – *Pereira, Fernando C. N., and Barbara J. Grosz* (eds). Artificial intelligence, vol. 63, nos 1-2.

Manning, Schütze, 1999 – *Christopher Manning and Hinrich Schütze* (1999). Foundations of Statistical Natural Language Processing, Cambridge, Massachusetts, USA. MIT Press.

Jurafsky, Martin, 2008 – *Daniel Jurafsky, James Martin* (2008). An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Second Edition. Prentice Hall.

Association – *Association for Computational Linguistics*. URL: https://www.aclweb.org/portal/.

Busa, 1990 – *R. Busa* (1990). Informatics and new philology, Computers and the Humanities, vol. 24, pp. 339-343.

Ospanova, 2014 – *B. R. Ospanova* (2014). Using Mathematical Methods and Information Technologies in Linguistics, Life Science Journal, 11(4s), pp. 373-375.

Klimova, 2009 – *Blanka, Klimova* (2009). Modern information technologies in teaching foreign languages, Bulletin of the Transilvania University of Braşov, 2(51), series IV.

Ghasemi et al., 2011 – *Babak Ghasemi, Masoud Hashemi, Simin Haghighi Bardine* (2011). The capabilities of computers for language learning, *Procedia-Social and Behavioral Sciences*, 28, published by Elsevier Ltd.

Chekharin, 2016 – *E. Chekharin* (2016). Information interaction in computational linguistics, *Slavic Forum*, 3(13), pp. 322-327 [in Russian].

Sinyagovska, Kryzhovska, 2020 – *I. Sinyagovska, V. Kryzhovska* (2020). Structural features and functions of computer terms in the sphere of information technologies, *Intellect. Speciality. Civilization*, 1(20), pp. 49-55 [in Ukrainian].

Petrasova, 2020 – *S. V. Petrasova* (2020). Modern information technologies in linguistics, Kharkiv: Panov A. M., 124 p. [in Ukrainian].

Moiseeva, 2017 – *I. Moiseeva* (2017). Quantitative linguistics and new information technologies, Orenburg: Orenburg State University, 101 p. [in Russian].

Prokopenko, Gartsunova, 2020 – *A. Prokopenko, L. Gartsunova* (2020). Fundamentals of Informatics and Applied Linguistics with a content module: copywriting, Sumy, 108 p. [in Ukrainian].

Potapova, 2007 – *R. K. Potapova*, (2007). Nanotechnologies and linguistics: Forecasting and perspectives of interaction, *Proc. Int. Sci.-Pract. Conf. "Nanotechnologies in Linguistics: Myth or Reality? Experience of Formation of Common Educational Space of SNG-Countries"*, Moscow [in Russian].

Guslyakova, 2016 – *A. V. Guslyakova* (2016). Information technology and linguistics of the XXI century, Moscow, 96 p. [in Russian].

Educational and methodical – *Educational and methodical complex of the discipline "Applied Linguistics"*. URL: http://fif.mdu.edu.ua/wp-content/uploads/2017/07/Prikladna-lingvistika-1-kurs.pdf.

English Wikipedia – *English Wikipedia*. URL: https://en.wikipedia.org/wiki/English_Wikipedia.

Computational Linguistics, 2014 – *Computational Linguistics* (2014). Stanford Encyclopedia of Philosophy, principal editor: Edward N. Zalta. URL: https://plato.stanford.edu/entries/computational-linguistics/.

Hobbs – *Jerry Hobbs*. Computers and Language. URL: https://www.linguisticsociety.org/content/computers-and-languages.

Encyclopedia – *Encyclopedia of Mathematics*. URL: https://encyclopediaofmath.org/wiki/Mathematical_linguistics.

Crystal, 2001 – *D. Crystal* (2001). Language and the Internet, Cambridge: Cambridge University Press.

Franchuk, 2019 – *N. P. Franchuk* (2019). Theoretical aspects of the study of linguistic informatics as an applied linguistic discipline, *Scientific journal of NPU named after Drahomanov*, Series 2, no 21, pp. 78-83 [in Ukrainian].

Kozerenko, 2006 – *E. Kozerenko* (2006). Linguistic aspects of informatics, *Sistemy i Sredstva Inform.*, spec. issue, pp. 88-111 [in Russian].

Yatsko, 2014 – *V. A. Yatsko* (2014). Computational linguistics or linguistic informatics? *Automatic Documentation and Mathematical Linguistics*, vol. 48, pp. 149-157.

Warschauer, Healey, 1998 – *Mark Warschauer, Deborah Healey* (1998). Computers and language learning: an overview, Lang. Teach., 31, Cambridge University Press, pp. 57-71.

Trinity College – *Trinity College*. Dublin, Computer Science Linguistics and a Language. URL: https://www.tcd.ie/courses/undergraduate/az/course.php?id=DUBMF-CSLA-3F09.

University of Groningen – *University of Groningen*, Language and Communication Technologies. URL: https://www.rug.nl/masters/language-and-communication-technologies-double-degree-programme/?lang=en.

Uppsala Universitet – *Uppsala Universitet*, Master's Programme in Language Technology 2021/2022. URL: https://www.uu.se/en/admissions/master/selma/program/?pKod=HSP2M&flik=0.

University of Washington – *University of Washington*, Master of Science in Computational Linguistics. URL: https://www.compling.uw.edu/about/.