20. TECHNOLOGY OF SYNERGY REVEALING IN TEACHING OF MATHEMATICS Prof. Dr. Eugeny Smirnov, Assoc. Prof. Dr. Artiom Uvarov, Senior Lecturer Nikita Smirnov, Yaroslavl State Pedagogical University named after Ushinsky, Russia .........................................................................................................................185

21. USE OF MULTIMEDIA TRAINING PRESENTATIONS IN TEACHING A FOREIGN LANGUAGE IN A NON-LINGUISTIC UNIVERSITY Prof. Valentin Zhigadlo, Assoc. Prof. Dr. Maria Odinokaya, PhD student Oksana Sheredekina, Smolny Institute of the Russian Academy of Education, Russia .................................197

22. VALUES OF THE SELF-DETERMINATION OF YOUNG PEOPLE (TEENAGERS) WHO ARE IN VARIOUS EDUCATIONAL SITUATIONS Prof. Antonida Lukina, Assoc. Prof. Olga Finogenova, Assoc. Prof. Zinaida Eflova, Assoc. Prof. Irina Peterson, Master Student Galina Reut, Siberian Federal University, Russia ..................................................................................................................205

Section HISTORY

23. THE EARLY PERIOD OF THE “DIGITAL REVOLUTION” FROM THE POINT OF VIEW OF INTELLECTUAL HISTORY Assoc. Prof. Aleksei Miloslavov, Assos. Prof. Olga Kuzmina, Saint-Petersburg State University, IFMO University, Russia .........................................................................................................................217

24. THE TRADE UNION MOVEMENT IN POLAND DURING 1918-1939 Dr. Adam Laska, Rzeszow University of Technology, Poland ......................................................................................................................227

Section LANGUAGE AND LINGUISTICS

25. BEHAVIORAL PATTERNS OF FARMERS IN RUSSIAN CULTURE (ON THE MATERIAL OF EVERYDAY TEXTS) Prof. Dr. Nailya Fattakhova, PhD in Philology Anastasiya Mubarakshina, Assoc. Prof. Nailya Faizullina, Kazan (Volga region) Federal University, Russia .........................................................................................................................239

26. CAREER GUIDANCE SERVICES IN EMPLOYER ENTITIES AND LIFE-LONG EDUCATION Assoc. Prof. Ing. Marta Matulčíková, PhD, Assoc. Prof. Dr. Daniela Breveniková, CSc, University of Economics in Bratislava, Slovakia ......245

27. CHARACTERISTIC FEATURES OF ACADEMIC LANGUAGE IN UNIVERSITY ENVIRONMENT WITH FOCUS ON THE GERMAN LANGUAGE (KEGA 029EU-4/2016) Assoc. Prof. Dr. Katarína Seresová, University of Economics, Slovakia .........................................................................................................................255

28. CODE SWITCHING AS A MANIPULATIVE STRATEGY IN RUSSIAN ISLAMIC DISCOURSE Assoc. Prof. Dr. Oksana I. Aleksandrova, Assoc. Prof. Dr. Svetlana A. Moskvicheva, Peoples Friendship University of Russia, Russia ..............263

29. DICTIONARIES ON SAMOYEDIC LANGUAGES AND LINGVODOC SOFTWARE SYSTEM FOR COLLABORATIVE WORK ON DICTIONARIES AND ONLINE PUBLISHING Dr. Julia Normanskaja, Institute of Linguistics of the Russian Academy of Sciences, Russia .........................................................................................................................273
30. **FUNCTIONAL PECULIARITIES OF THE VERB TÄJT/TÄJT ‘HAVE’ IN NORTHERN KHANTY DIALECTS (KAZYM, SHURYSHKAR AND PRIURAL)**
Assoc. Prof. Victoria Vorobeva, Prof. Dr. Irina Novitskaya, Tomsk State University, Tomsk Polytechnic University, Russia

31. **GRAPHIC FEATURES OF THE ARCHIVAL MANSI LANGUAGE DICTIONARY, COMPILED BY ARCHPRIEST SIMEON CHERKALOV (1783)**
PhD Student Natalia Koseliuk, Dr. Yu.V. Normanskaya, National Research Tomsk State University, Russia

32. **INTERNET ARCHIVING: THE USE IN DISCOURSE STUDIES**
Dr. Olga Solopova, Assoc. Prof. Dr. Maria Saltykova, Assoc. Prof. Dr. Maria Voroshilova, South Ural State University (National Research University), Russia

33. **PHILOSOPHICAL DISCOURSE OF F. BACON AS THE REFLECTION OF PHILOSOPHICAL TRADITIONS OF RENAISSANCE EPOCH**
Prof. Zinaida Fomina, Assoc. Prof. Irina Lavrinenko, Assoc. Prof. Victoria Kozlova, Voronezh State Technical University, Russia

34. **POETICS OF CHRONOTOPE IN NOVELS BY NIKOLAY LUGINOV**
Doctor Okorkova V.B., Assoc. Prof. Permyakova T.N., Assoc. Prof. Saburova N.V., Sakha Literature Department, NEFU Institute of Languages and Cultures of the Peoples of the Northeast, Russia

35. **RECONSTRUCTION OF THE 14TH AND 15TH PROTO-SAMOYEDIC VOWELS**
Prof. Dr. Julia Normanskaja, National Research Tomsk State University, Moscow Institute of Linguistic RAS, Russia

36. **THE ROLE OF THE LANGUAGE IN MIGRATIONS AT THE TURN OF THE CENTURY IN CROATIA**
Assoc. Prof. Dr. Emina Berbic Kolar, Dr. Igor Marko Gligoric, University of Osijek, Faculty of Education, Croatia

Prof. Dr. Fanuza Nurieva, Milyausha Sungatullina, Assoc. Prof. Margarita Petrova, Kazan Federal University, Russia

38. **WILLIAM HAZLITT’S ESSAYS ON WRITING AND STYLE AS A SOURCE FOR ACADEMIC WRITING**
Assoc. Prof. Dr. O. Anossova, Peoples’ Friendship University of Russia, Russia

Section PHILOSOPHY

39. **CULTURE AND CIVILIZATION: FUNCTIONAL AND METHODOLOGICAL ASPECTS**
As.Pr.Evgeniya Kuznetsova, As.Pr.Olga Goryacheva, As.Pr. Gulchachak Patenko, University of Management “TISBI”, Russia

40. **DEATH IMAGES IN RUSSIAN YOUTH SUBCULTURES**
Ph. D., Associate Professor, Fedorova Marina, Linguistic University of Nizhny Novgorod, Russia
INTERNET ARCHIVING: THE USE IN DISCOURSE STUDIES

Dr. Olga Solopova
Assoc. Prof. Dr. Maria Saltykova
Assoc. Prof. Dr. Maria Voroshilova

1South Ural State University (National Research University), Russia
2Ural State Pedagogical University, Russia

ABSTRACT

Digitization of archival materials has become extremely popular in the Internet data storage era. Nowadays on-line archives contain large-scale digital resources and provide access to data from almost anywhere without visiting the archives in person. On-line archival records are a unique source of information as they preserve historic, often landmark, materials (newspaper, magazines, photographs, letters, reports, audiovisual recordings, etc.) and make them available for use via the Internet. The authors argue that Internet archival records can be effectively used in the research on retrospective and diachronic studies in discourse analysis. On-line digital collections contain original sources and relevant artifacts from federal, state and local institutions with physical archives of different countries. The authors focus on a) the types of on-line digital collections, b) the methods used to study archival materials c) the finding aids that help in identifying relevant documents and records characteristics that convey important clues to identify them, d) the benefits of using Internet archival collections to better understand the time-sensitive nature of discourse. The authors exemplify the key points discussed in the paper with the results of their recent research on World War II period. The paper is of interest to a wide range of experts in linguistics, history, sociology, political linguistics, political science. As digital content is often full-text searchable and quite easy to work with, it can be also used in teaching and studying linguistic disciplines (historical linguistics, corpus linguistics, discourse analysis, forensic language expertise, etc.).

Keywords: archival materials, on-line digital collections, linguistics, discourse studies, World War II.

INTRODUCTION

Nowadays all of us have got used to a new method of getting and providing information – through electronic media. The Internet has become “the agora for research, teaching, expression, publication, and communication” [1]. These media include various resources that can be used both for education and research. One of them is on-line archives of different physical cultural institutions: museums, libraries, archives, governmental agencies, institutions of higher educations, etc. Digital libraries offer free universal and long-term access to collections of cultural artifacts. Internet archives preserve textual documents, photographs, maps, tape recordings, graphic records, line drawings, artistic illustrations, and other similar records. Digitization of archival materials helps to protect rare, unique, often fragile, historical documents and art objects transferring them to a digital medium.
and delivering to end-users. Newly designed user interfaces, full-text searching, cross-collection indexing allow scholars to study corpora of artifacts recording the history of humankind. “Corpus linguistics is maturing methodologically and the range of languages addressed by corpus linguists is growing annually” [2]. Thus, corpus-based research (comprising both quantitative and qualitative findings) is an efficient way to analyze language in use, or discourse, from synchronic, retrospective and diachronic viewpoints. (The authors illustrate the key points discussed in the paper with the samples of their recent corpus-based research. The illustrative examples cited here were taken from different on-line digital collections).

**TYPES OF CORPORA**

Corpora provide data, so a corpus-based study aims to gather findings that test a certain idea or model, whatever its scope [3]. The researcher can use a corpus as an empirical base for illustrating and verifying hypotheses or make it the starting-point for a certain study. The last-mentioned approach is denoted as corpus-driven [4], [5]. “If the approach is corpus-driven it is inductive (bottom-up) since the linguist aims at arriving at a theory starting from the empirical data” [6]. Otherwise the researcher follows a deductive approach in order to validate or refute the theory that is applied to the corpus data. Though, both approaches are commonly combined. In order to examine archival materials the lead question that needs to be considered is which corpus to explore. From this point there is a large variety of corpus types for use via the Internet.

A monolingual corpus contains texts in one language only and is the most frequent type of corpora. It is used by a wide range of users for various tasks: checking the correct usage of a word or looking up the most natural word combinations, identifying frequent patterns or new trends in language, etc. For example, the Tehran Monolingual Corpus, a large-scale Persian monolingual corpus, which comprises more than 250 million words.

A multilingual corpus contains texts in more than one language and is very similar to a parallel corpus that consists of two monolingual corpora: these terms are often used interchangeably. Multilingual corpora comprise texts in several languages which are all translations of the same text and are aligned in the same way as parallel corpora, so researchers can take a look into the behavior of the same word or phrase in different languages. An example is the Enabling Minority Language Engineering (EMILLE) corpus [7]. Corpora of this kind are particularly useful in translation and contrastive studies, gain new insights, as compared to monolingual corpora, highlight language-specific, typological, or cultural features and can be useful for lexicography.

Learner corpora collect the language output produced by learners of the second language. This type of corpora helps to build a profile of learner language, particularly in terms of error analysis or for fact-finding what words, phrases, parts-of-speech are over- or under-used by learners, compared to native speakers. International Corpus of Learner English (ICLE), Cambridge Learner Corpus part of the Cambridge International Corpus (CIC), Longman Learners’ Corpus, Standard Speaking Test (SST) Corpus, Chinese Learner English Corpus (CLEC), for example.
A monitor (or dynamic) corpus is a type of diachronic corpora that may continue to grow with new texts added over time. It is currently used to track changes across different periods of time. The Global English Monitor Corpus, for example, collects newspapers in English and tracks language use and semantic change across the English language discourses in Britain, the United States, Australia, Pakistan and South Africa.

A synchronic corpus represents the language as a data base in which all of the texts have been collected from roughly the same time period, allowing a ‘snapshot’ of language use at a particular point in time [7]. A typical synchronic corpus is the International Corpus of English (ICE) specifically designed for the synchronic study of ‘world Englishes’. Its data base consists of a collection of twenty corpora of one million words each, and it is composed of written and spoken English produced during 1990–1994 in countries or regions where English is either the first language or an official language.

One of productive ways to explore language variation is from a diachronic perspective using diachronic corpora that look at changes across a timeframe. That data base contains texts covering a wide range of time periods and is used by researches to track and study linguistic changes within it. Take the British Newspaper Archive [8], for example. As this on-line collection is based on the physical archive of the British Library, it contains highly representative texts. This digitization project is diachronic as it includes 25,274,255 pages dating from the 1700s. The British Library’s digitization of their historic newspapers, containing most of the runs of newspapers published in the UK since 1800, has turned this corpus into one of the finest in the world. They have even been able to scan some of the rarest and most fragile newspapers in the collection. The British Newspaper Archive is a monitor corpus, as it continues to grow with new texts added over time. Now they are adding a wealth of material from the 20th century, right up to the 1950s. This corpus is monolingual as it includes newspapers in one language – English. The scale of the newspaper publishing industry is enormous:

- with newspapers from England, Scotland, Ireland, and Wales,
- with national, local, provincial newspapers and journals (newspapers aimed for county circulation only),
- with many cities and towns publishing several newspapers simultaneously,
- with newspapers aimed at distinct audiences depending on social status, geographical location and political affiliation.
- with different material types: news articles about global and national events, as well as issues of local and regional importance; letters to the editor written by the newspaper’s readers, including illuminating contemporary debates, aspirations and anxieties; obituaries that are sources of contemporary information on the lives of notable individuals; advertisements; photographs; graphics; maps; editorial cartoons. The corpus doesn’t have any linguistic annotation, it contains only extra-linguistic tags: the title of the newspaper, the location, the date, the number of the issue, the number of the page the text is on. Thus, the British Newspaper Archive provides an unrivalled picture of historical life spanning any retrospective period chosen for discourse analysis.
There are many other types of corpora for studying language and discourse variation via the Internet: a reference or target corpus (a corpus whose data are used to compare with those of other corpora, usually through statistical data analysis), tagged and raw corpora (with or without annotation), specialized corpora (limited to one or more subject areas, domains, topics, etc.), multimedia corpora (containing audio or visual materials or other type of multimedia content), parsed, spoken and national corpora, etc. Each corpus can come in many shapes and the choice of the ‘right’ corpus depends on the aims of the study: it is equally important to relate the results from the corpus analysis to the nature of the corpus.

METHODS

Working with corpora in discourse studies it is essential to carefully choose the corpus and set criteria that will ensure that the findings will be representative. For example, the British Newspaper Archive allows the investigator to search hundreds of millions of articles by a keyword, a name, or a title and watch results appear in an instant. The user interface comprises the finding aids that help in identifying relevant documents. There are two ways to get to the particular documents you are interested in: searching or browsing. The ‘Advanced Search’ box combines several powerful options and lets the investigator type in specific keywords, fill in date ranges, regions, titles, article types. The British Newspaper Archive uses what is known as a fuzzy search including the exact spelling of the word that the investigator enters and any related word forms. The ‘Exact Search’ checkbox excludes any related word variants from the search results. Browsing lets the researcher find a specific newspaper or the exact edition he’s looking for by its title or by its place of publication.

As the aim of our recent research was to find out conceptual metaphorical images associated with the future of the Soviet Union in British discourse of World War II period, we narrowed down our selection by using the ‘Advanced Search’ box. The options we selected included the search terms: Russia, USSR, Soviet, future; the publication date: from 01 September 1939 to 02 September 1945 (the period of World War II); with all types of publication places and titles included; with ‘articles’ and ‘illustrated articles’ chosen among article types; with ‘results’ sorted by relevance. The illustrative corpus comprised 18810 documents that could be further sorted by a specific year, by each country of the UK, by the article type (Fig. 1), as well as by regions, counties and places of publication.
### Figure 1. Search results (sorted by a specific year, article type, country)

All documents matching the search query were shown in the results panel. When ordered by relevance, the search results that contained the exact matches appeared first (Fig. 2). Once the Viewer appeared, we were able to move through the newspaper pages using a mixture of zoom, scrolling and grab tools. The Viewer in the British Newspaper Archive contains both print and download image options.

**Sweeping Changes Made in Russia**  
Wide Political Soviet Republics

> ...CHANGES MADE IN RUSSIA Wide Political Soviet Republics FAR-REACHING changes extending the powers of the 16 Republics of the were introduced by Mr. Molotov, Soviet Commissar for Foreign Affairs, in a speech yesterday to the Supreme Soviet assembled in Moscow...

**Published:** Wednesday 02 February 1944  
**Newspaper:** Yorkshire Post and Leeds Intelligencer  
**County:** West Yorkshire, England  
**Type:** Illustrated | Words: 5632 | Pages: 1-2 | Tags: none

### Figure 2. Fragment of search results

After retrieving the information from the corpus we went beyond the quantitative patterns and used the methods of metaphorical modeling, component analysis and contextual analysis, cognitive and discourse analysis and proposed interpretations explaining why the patterns existed.

The method of metaphorical modeling comprised evaluating the productivity of each metaphor fixed while analyzing the data obtained (Table 1), describing its frame structure, characterizing the ability of metaphors to depict negative and positive images of the future of the USSR.
Table 1. Retrospective system of metaphors functioning in British discourse (the future of the USSR being a target domain)

<table>
<thead>
<tr>
<th>№</th>
<th>THE FUTURE OF THE USSR IS</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INANIMATE NATURE</td>
<td>14,2%</td>
</tr>
<tr>
<td>2</td>
<td>PATH</td>
<td>12,5%</td>
</tr>
<tr>
<td>3</td>
<td>HOME</td>
<td>12,4%</td>
</tr>
<tr>
<td>4</td>
<td>FAMILY RELATIONSHIPS</td>
<td>11,4%</td>
</tr>
<tr>
<td>5</td>
<td>BUILDING</td>
<td>9,3%</td>
</tr>
<tr>
<td>6</td>
<td>ORGANISM</td>
<td>8,9%</td>
</tr>
<tr>
<td>7</td>
<td>MECHANISM</td>
<td>8,0%</td>
</tr>
<tr>
<td>8</td>
<td>FAUNA</td>
<td>4,2%</td>
</tr>
<tr>
<td>9</td>
<td>WAR</td>
<td>3,8%</td>
</tr>
<tr>
<td>10</td>
<td>TEACHING AND LEARNING</td>
<td>2,8%</td>
</tr>
<tr>
<td>11</td>
<td>MARKET</td>
<td>2,5%</td>
</tr>
<tr>
<td>12</td>
<td>DESEASE</td>
<td>2,5%</td>
</tr>
<tr>
<td>13</td>
<td>MORNACHY</td>
<td>2,1%</td>
</tr>
<tr>
<td>14</td>
<td>FLORA</td>
<td>1,6%</td>
</tr>
<tr>
<td>15</td>
<td>SPORT</td>
<td>1,2%</td>
</tr>
<tr>
<td>16</td>
<td>GAME</td>
<td>0,8%</td>
</tr>
<tr>
<td>17</td>
<td>THEATER</td>
<td>0,7%</td>
</tr>
<tr>
<td>18</td>
<td>INSTRUMENT</td>
<td>0,5%</td>
</tr>
<tr>
<td>19</td>
<td>CRIME</td>
<td>0,5%</td>
</tr>
</tbody>
</table>

"The corpus linguistics largely supports the cognitive theory of metaphor" [9] that was first extensively explored by G. Lakoff and M. Johnson [10]. The methods of component analysis and contextual analysis describe the application of the meaning to the word and its subsequent use in the language and help to consider all the circumstances in the emergence of each metaphor, its paradigmatic and syntagmatic relationships with other linguistic means. The method of cognitive discourse analysis helps to assess the metaphors and the texts in their historical, cultural or social context [11], [12], [13].
ADVANTAGES

There are numerous practical advantages of corpus techniques in using online archives: "due to their speed and the sheer quantity of occurrences they clearly allow research to be undertaken that would otherwise never have been completed because of lack of time and human resources" [3]. One of the apparent advantages of using corpora is a high possibility of making an unbiased data analysis retrieving information remotely. Not only the researcher's linguistic competence but also a great amount of authentic language material allows the researcher to pursue a more objective way of studying language in use [6]. Nowadays the availability of much larger corpora makes it possible to evaluate a lot of data at once. "We need a lot of text so that there might always be a sufficient residue of useful examples" [14]. It is another advantage to test diversified corpora that comprise material taken from different sources such as online national libraries containing texts with a high degree of representativeness. "The language looks rather different when you look at a lot of it at once" [14].

'Technical' benefits of using Internet archival collections consist in the possibility of combining several different databases. A good example is Gallica [15], the on-line library of National Library of France. Its data base provides free access for all kinds of documents: books, magazines, newspapers, photographs, caricatures, posters, maps, manuscripts, scores, audio and visual recordings, book-miniatures, etc. The large number of observable data collections, the possibility of studying on wide-ranging diachronies and of 'zooming' different periods of language evolution will definitely result in a fresh take on language variation and its periodization. For example, due to new means of analysis and use of large corpora such on-line collections as 'Very Old French' or 'Pre-Classical French' have emerged [3]. Working with corpora in discourse studies can help to bring to the fore both differences and specific features within subsystems regarding particular vocabulary, phraseology, syntax, etc.

Another important benefit to be emphasized is opposed to the usual linear reading of a single text. It's the 'vertical' reading of texts via concordances, also referred to as keywords in the context (KWIC) [3], [7]. A concordance is "a list of all of the occurrences of a particular search term in a corpus, presented within the context in which they occur" [7]. The use of concordances makes it possible to scroll the whole document at once through 'clipping' contexts by such means as highlighted keywords and to reveal fragments that would otherwise have gone unnoticed. Thus, the researcher can extract all necessary data in an instant launching automated distributional program based on various contextual parameters set beforehand (Figure 3).
These technical benefits have led to a true renewal of discourse studies and historical linguistics in general.

Besides, being full-text searchable and quite easy to work with, digital content can be also used in teaching and studying different linguistic disciplines (historical linguistics, corpus linguistics, discourse analysis, forensic language expertise, etc.). That’s due to the facts that the new target audience of on-line archives is the general education market and the new method of providing information for the younger generation is through electronic media, which means that on-line libraries and archives tend to become primary information providers.

**CONCLUSION**

In our Internet data storage era digitization of archival records is fast developing. On-line archives and libraries contain large-scale original resources. Internet archiving serves local, federal, national and international needs. Various types of on-line digital collections (monolingual, multilingual, monitor, diachronic, synchronic, target, learner, etc.) can be used for different purposes in researching, teaching and learning. Digital archives of physical cultural and educational institutions offer their users unique advantages: consulting libraries and archives remotely, examining a large amount of data within a short period of time, full-text searching, cross-collection indexing, special finding aids that help in identifying relevant documents, ‘vertical reading’, etc. Thus, these digital conversion projects inspire new scholarly work, while the access to unique or special collections’ material makes discourse research (both synchronic and diachronic) much easier. Any corpus-based study comprises a large number of related methods used by scholars whose aim is to explain and to exemplify quantitative and qualitative patterns of the discourse phenomena under analysis. The findings of corpus-based studies can be used in teaching and learning a number of linguistic disciplines.
ACKNOWLEDGEMENTS

The research is financially supported by the Russian Scientific Foundation, project No. №16-18-02102.

REFERENCES


