

# Examination of Trends in Education with the Google Books Ngram Viewer

Vladimir M. Moskovkin, Tatyana V. Saprykina, Elena V. Pupynina, Vladimir A. Belenko, Irina A. Shumakova

**Abstract:** *In the modern world there has been change in education environment and education technologies, which has led to new concepts and tendencies in training and education. The purpose of this research is to explore evolution of terminology in distance, open, online education and learning. The change in use of the terminology is analysed through the Google Books Ngram Viewer that allows to define frequency of the use of the terms in six publications corpora for the certain period. The research identifies tendencies of the use of the terms in the sphere of education in different language groups compared and reveals interrelation between the change in the use of the terminology and economic, social, political and technological factors.*

**Keywords :** *education, distance education, distance learning, remote courses, e-learning, online courses, online education, the Google Books Ngram Viewer.*

## I. INTRODUCTION

In the modern world there has been change in education environment, education technologies, which has led to new concepts and tendencies in training and education. Traditional education implies ongoing interaction between a teacher and a student in an educational institution, it involves lectures and seminar classes with further assessment. However, development of society, infrastructure, elaboration and use of new means of communication give rise to new approaches to teaching. Researchers in many countries investigate development of the modern education technologies. One of them is distance learning and the use of the Internet technologies in education.

As life becomes more intensive people become more interested in distance way of getting education. In his book “Foundations of distance education”, Keegan writes, “The generation from 1970 to 2000 is witnessing a development of the whole field of distance education which paralleled the success and achievements of the Open University. There has been a remarkable change in the quality, the quantity, the status and the influence of distance education provision. This was linked to a world-wide shift from private to public provision.” [1: 3].

**Revised Manuscript Received on July 22, 2019.**

\* Correspondence Author

**Vladimir M. Moskovkin**, Belgorod State University, 85 Pobeda St, Belgorod, 308015, Russia.

**Tatyana V. Saprykina**, Belgorod State University, 85 Pobeda St, Belgorod, 308015, Russia.

**Elena V. Pupynina**, Belgorod State University, 85 Pobeda St, Belgorod, 308015, Russia.

**Vladimir A. Belenko**, Belgorod State University, 85 Pobeda St, Belgorod, 308015, Russia.

**Irina A. Shumakova**, Belgorod State University, 85 Pobeda St, Belgorod, 308015, Russia.

One can assume that development of the Internet became the next stage in advancement of distance education. In the book “Technology, e-learning and distance education” Bates shows implementation and development of the Internet technologies in teaching, distance education and writes about the need to restructure educational institutions: “Distance education has resulted in major changes to the organization of educational provision. The obvious difference is students are no longer required to attend a campus at regular periods. As a result, distance education has required quite different organizational structures from those of conventional educational institutions. Furthermore, as technology changes, so does the need to reorganize institutions to lever the benefits of new technology.” [2: 17].

Factors to be considered when planning and organizing distance learning have been revealed. For example, apart from the need to change organizational structure, Levy writes about the necessity to take into consideration such factors as planning, staff and student training and support, copyright and intellectual property [3].

The change in learning technologies influences development of concepts. Words and phrases statistics and analysis provide quantitative data on frequency and time periods of their use. Baldassarre notes that in a few seconds the Ngram Viewer service graphs the change in the use of words over time [4].

The issues of the change in terminology in education system are dealt with in the article by Silber-Varod, Eshet-Alkalai, & Geri “Analysing the Discourse of Chais Conferences for the Study of Innovation and Learning Technologies via a Data-Driven Approach”. The authors use the Ngram Viewer to find out that learning is becoming more connected to the Internet. The authors conducted their research on the basis of the corpora in two languages: Hebrew and English. They found out that “the most frequent keyword in the examined corpus was “collaborative learning” (40 instances). This was followed by “distance learning” and “E-learning” in 25 and 23 of the articles respectively” [5].

## II. METHODS

To perform the complete research of the evolution of terms connected to distance, online and open education as well as emergence and development of certain concepts in this area we use the Google Books Ngram Viewer to analyse their frequency. The Ngram Viewer tool was developed for culturologists, linguists and other researchers in humanities in 2010 by a team from Harvard University in collaboration with Google and became widely used [6, 7, 8].

The Russian terms were taken as the basis for the research: ДИСТАНЦИОННОЕ



## Examination of trends in education with the Google Books Ngram Viewer

образование (remote education / distance education), дистанционное обучение (distance learning), открытое обучение (open training), открытое образование (open education), дистанционные курсы (distance courses), онлайн-курсы (online courses), онлайн-образование (online education), онлайн-обучение (online training), онлайн-образование (online education), онлайн-обучение (online training), интернет-образование (Internet education), интернет-курсы (Internet courses),

онлайн-университет (online University), интернет-университет (Internet University), онлайн-университеты (online Universities) (totally 15 terms).

The analysis of frequency of the terms and their further research were conducted on the basis of corpora in six languages: Russian, German, English, Spanish, Italian, and French (Table 1).

Table 1: Terms connected to distance, online and open education in six languages corpora

Groups	Russian	German	English	Spanish	Italian	French
1. General	Дистанционное образование	Fernunterricht	Remote education Distance education	Educación a distancia	Formazione a distanza	Enseignement à distance
	Дистанционное обучение	Fernstudium	Distance learning	Aprendizaje a distancia	Apprendimento a distanza	Enseignement à distance
	Открытое обучение	Training öffnen Offene Ausbildung	Open training	Entrenamiento abierto	Formazione aperta	Entraînement ouvert Formation ouverte
	Открытое образование	Offene Bildung	Open Education	Educación abierta	Open Education	éducation ouverte
2. Specific	Дистанционные курсы	Fernkurse	Distance Courses	Cursos a distancia	Corsi di distanza	Cours à distance
	Онлайн-курсы	Online-Kurse	Online Courses	Cursos en línea	Corsi online	Cours en ligne
	Онлайн-образование	Online-Bildung	Online Education	Educación en línea	Istruzione online	Éducation en ligne
	Онлайн-обучение	Online-Schulung	Online Training	Entrenamiento en línea	Formazione online	Formation en ligne
	Онлайн-образование	Online-Training	Online training	Capacitación en línea Entrenamiento en línea	Formazione online	Formation en ligne Éducation ouverte
	Интернет-образование	Internet-Bildung	Internet education	Educación en internet	Internet educazione	Éducation à Internet
	Интернет-курсы	Internetkurse	Internet Courses	Cursos de internet	Corsi di Internet	Cours Internet
	Онлайн-университет	Online Universität	Online University	Universidad en línea	Università Online	Université en ligne
	Интернет-университет	Internet Universität	Internet University	Universidad de Internet	Università di Internet	Université Internet
	Онлайн-университеты	Online-Universitäten	Online universities	Universidades en línea	Università online	Universités en ligne

To enhance reliability of the results the analysis with the help of the Google Books Ngram Viewer was performed many times during the period from April 2016 to December 2017. It should be noted that corpus in a certain language does not only include the number of works originally published in that language. Translated works are rather frequently found within corpora. Quantitative data also includes phrases found in References sections of publications. If this occurs, such phrases may only indirectly be connected to the content of a publication or may convey aspects other than those being studied.

It is also important to mention that the Ngram Viewer tool makes it possible to search for no more than twelve terms simultaneously for direct comparison. Furthermore, there are language specific aspects. For example, Russian is an inflected language and words with different case endings are treated by the Ngram Viewer tool as separate words grouped into different trend lines.

To raise the objectivity of the results the search was conducted with the case insensitive option selected. It allows to trace and group as identical terms used at the beginning of sentences (with capital letters)



and within sentences (with small letters). If the corpus in a certain language contains terms with both upper-case and lower-case characters the Ngram Viewer graphs have mark “All” near such terms, which was also shown by Michel, J.-B. et al. [6].

### III. RESULTS

Frequency of occurrence of all the terms under study in the Russian corpus is shown in Figure 1. Within the range of term frequency change under question only 6 out of 15 terms are visible. Further in the article apart from the word “occurrence” such words as “incidence”, “mentioning”, “usage”, “use” will appear.



Figure 1. Frequency of all the terms under study in the Russian books corpus

Figure 1 demonstrates that the use of the phrases “distance learning” and “distance education” prevails in the corpus. The proportion of these terms is so big that it does not allow to trace the tendency of change of such phrases as “distance courses”, “open training”, “Internet education”.

The analysis of all the terms with Ngram Viewer leads to the unambiguous conclusion that the most frequent phrases are the ones with more general meaning such as “distance learning” and “distance education”. This trend is observed in corpora in all the six languages.

Taking into account various interpretations of concepts in education domain and due to quantitative limitations of queries in Ngram Viewer tool we divided 15 terms mentioned earlier into two groups (table 1) for the purpose of more profound research. It is reasonable to start with general terms: remote / distance education, distance learning, open training, open education. In this group of terms the first mentioning of the phrases “open training” and “open education” traces back to mid 1920s (Fig. 1).

It is worth noting that due to specific characteristics of the Russian language the phrase “открытое образование” (open education) may refer to the spheres of social life other than education. Thus, in the book “Mikhail Aleksandrovich Bakunin: yego zhizn i deyatelnost (1814-1876): Raskol v Internatsionale” (“Mikhail Aleksandrovich Bakunin: his life and work (1814-1876): The split in the International”) by Steklov this term is found in the following fragment: “... the first open social revolutionary word was formed... Then the newspaper published a series of his articles, out of which the most remarkable articles were the ones about... integral education” [9: 354]. This book appeared in the sources where the phrase “open education” occurred despite the fact that the words “open” and “education” are found in different paragraphs on the same page.

In Bolshaya Sovetskaya Entsiklopediya (The Great Soviet Encyclopaedia) (Volume 60) the same phrase (открытое образование) is homonymous to “open establishment”: “Open establishment of public unions and political groups, a number of government statements about the revolutionary

movement misled the censorship, deprived it of strictness...” [10]. In both sources cited above the phrase “открытое образование” occurs in different senses meaning establishment of a political alliance in the second case.

In the context of education this phrase occurs in late 1940s. For example it is mentioned in the journal Vestnik Vyssei Shkoly (Review of the Higher School): “The need for special transport education in our country arose quite far back in the past... Initially it was an open educational institution.” [11].

On the whole exponential growth of frequency of the phrases “distance learning” and “distance education” in the books corpus was observed in 1990s. Popularity of the terms started to decline in 2003-2004 (Fig. 1), which might be due to the fact that they were substituted for by new terms.

The phrase “distance education” was first used in Russia in mid 1990s. In 1995 the dictionary “Obrazovaniye vzroslykh: mezhdistsiplinaryny slovar terminologii” (Education of adults: interdisciplinary dictionary of terminology) by Onushkin & Ogaryov was published. It defined the term “distance education” in the following way: “Distance education – international term that is sometimes translated as remote education...” [12]. The same year was marked by adopting the Strategy of creating and developing the common network of DE (distance education) in Russia. Since 1993 Russian Independent University (now Russian New University) published the journal “Otkrytoe obrazovaniye” (Open Education) for two years; in 1996 Moscow Institute of Economics and Statistics started publishing the journal “Distantsionnoye obrazovaniye” (Distance Education), which has been published under the title “Otkrytoe obrazovaniye” (Open Education) by Plekhanov Russian University of Economics since 2000. Modern Humanitarian University has been publishing the journal “Distance and Virtual Learning” since 1999. These events correlate with growth in frequency of the respective terms, which is shown in Fig. 1.

In the German corpus the most frequent phrases are distance learning (Fernstudium) and distance education (Fernunterricht) (Fig. 2).



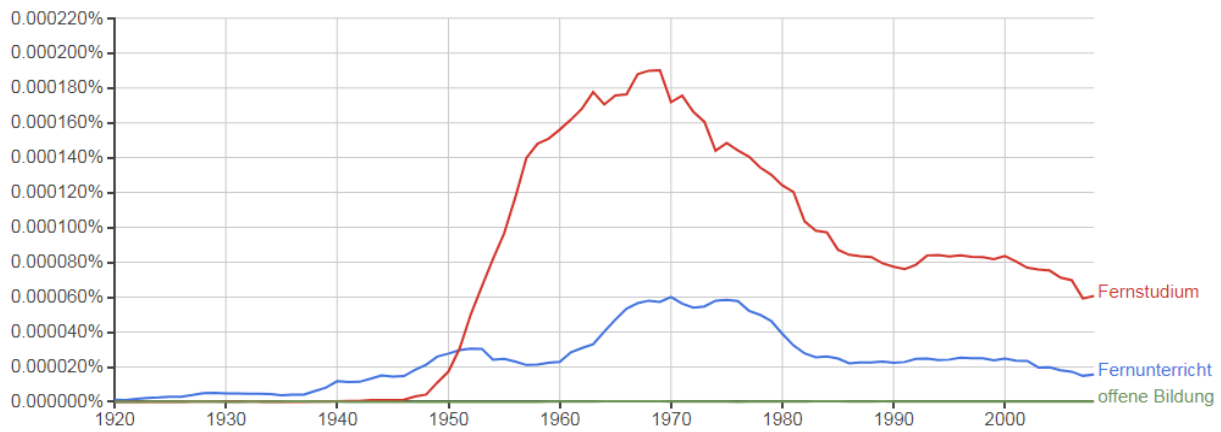


Figure 2. Frequency of the terms of the first group in the German corpus

Frequency of the terms peaks in the period of 1965-1970 (Fig 2). The reason is publication of reference books and a number of books such as «Zwei Jahrzehnte Bildungspolitik in der Sowjetzone Deutschlands, Dokumento» (1966) and «Tübinger Beiträge zum Fernstudium» (1970), where the terms are mentioned on pages 46 and 86 respectively, with the phrases occurring both in the text and in references section. It

is important to note that learning does not always refer to the provision of education in its traditional sense. In some sources open learning means sports training outdoors.

In the English corpus the terminology under consideration occurred a little later and the number of publications about distance education peaked in 2002-2005 (Fig. 3).

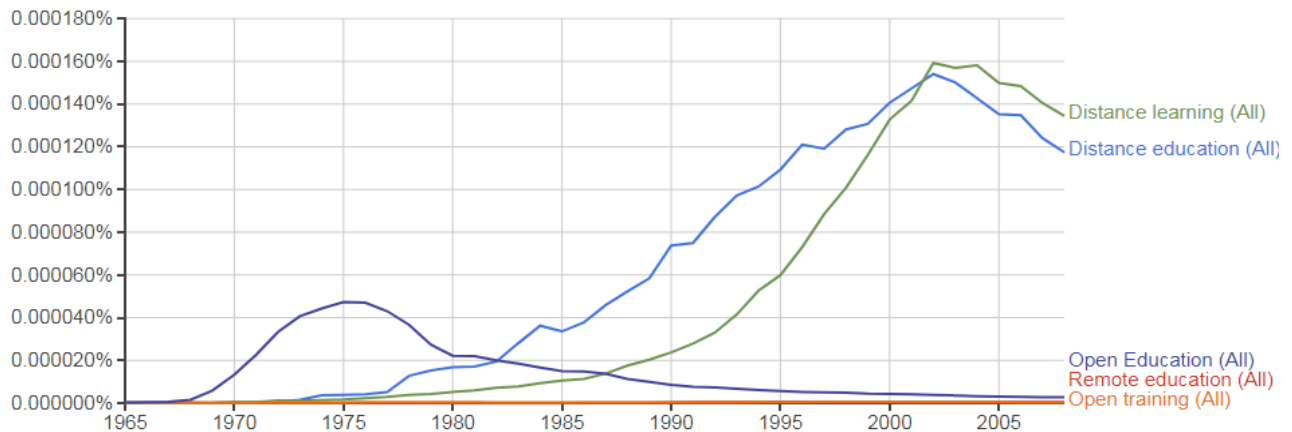


Figure 3. Frequency of the terms of the first group in the English corpus

However, open education attracted researchers' attention in late 1960. In 1975 the book "The Philosophy of open education" by Nyberg came out. That is why maximum frequency of the term "open education" is observed at this time. The book is a major work about how the philosophy of the concept "distance education" formed. In the book, the author states that open education is "that form of educational practice which is characteristically regulated by the following rules:

- (1) Students are to pursue educational activities of their own choosing;
- (2) Teachers are to create an environment rich in educational possibilities;
- (3) Teachers are to give a student individualized instruction based on what he/she is interested in, but they are also to guide the student along educationally worthwhile lines;

(4) Teachers are to respect students. The following count as exhibiting respect for the student:

- (a) the student is granted considerable freedom; he/she is, for the most part, autonomous,
- (b) the student's interests and ideas are considered to be important and he/she receives individual instruction and guidance based on his/her interests,
- (c) there is considerable interaction between teacher and student; they are considered to be equal in some sense,
- (d) students are rarely commanded; uses of authority are minimized,
- (e) students' feelings are to be taken seriously" [13].

In the Spanish corpus the concept of distance education occurs in early 1970s (Fig. 4).

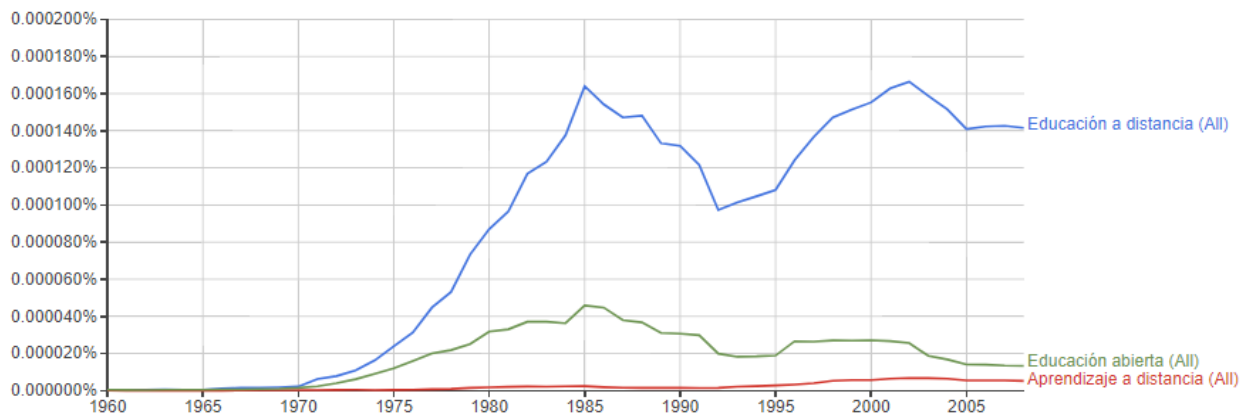


Figure 4. Frequency of the terms of the first group in the Spanish corpus

Many Spanish publications in Ngram Viewer service are stored in the USA. For example, Arrien «Los retos de la «educación a distancia»: I Seminario-Taller de Teleducación Universitaria de FUPAC» [14] and Educación [15] is held in University of Pennsylvania library, Pintado «La reforma educativa de la Segunda República Española: primer bienio» [16] is stored in University of California library.

Considerable number of publications where the terms under question occur is not directly connected to the description of the system of education. These terms appear in

bibliographical entries, abstracts, information about the authors, publications on other issues, etc.

However, there are some works in Spanish that are of interest to our research. For example, Instituto Colombiano para el Fomento de la Educación Superior (1986) published “Metodología y estrategias de la educación superior abierta y a distancia: nivel introductorio” (“Methodology and strategy of open and distance education”) [17].

The Italian and French corpora reflect the same trend as the Spanish one (Fig. 5, 6).

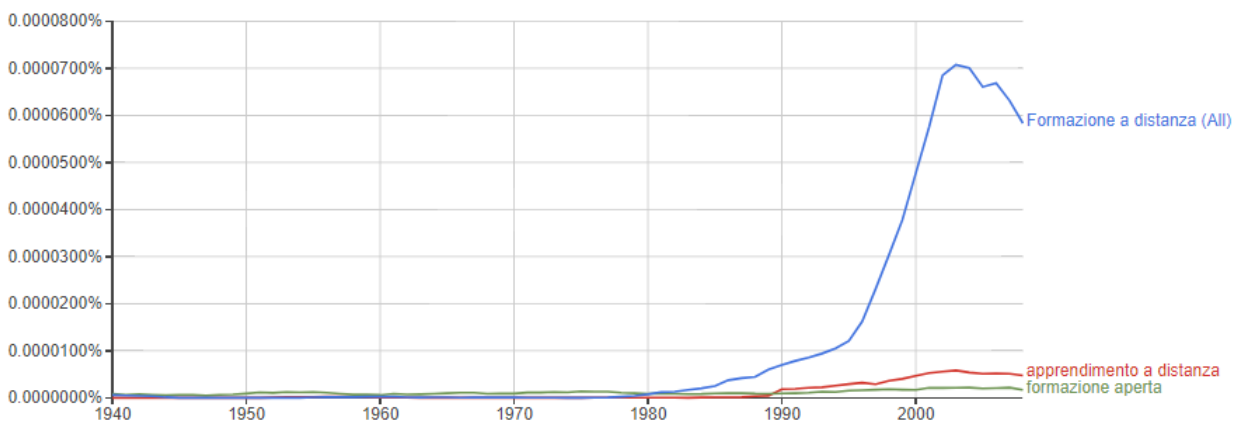


Figure 5. Frequency of the terms of the first group in the Italian corpus

The books corpora in Italian show that the term “distance education” is prevailing. Small increase in frequency of this term is seen from 1980 to 1995, after this period there is significant growth up to 2004. A considerable number of works were published during that time. Some of them are Merlo “Formazione e lavoro femminile: il contributo del fondo sociale europeo : politiche ed esperienze” [18],

Quagliata “Pratiche di didattica costruttivista in aula e nella rete” [19]. The content of the publications of that period presents both the sphere of education and improving computer competence (distance learning of computer programs).

## Examination of trends in education with the Google Books Ngram Viewer

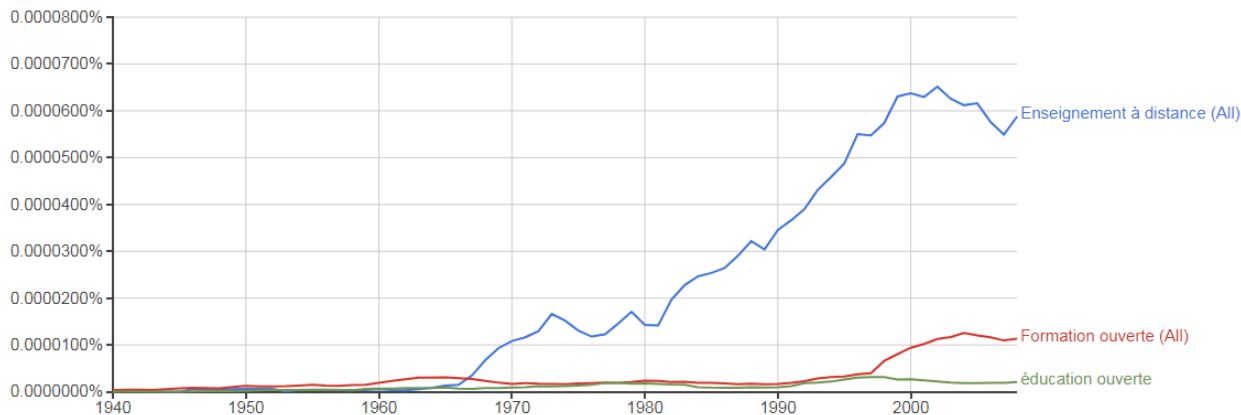


Figure 6. Frequency of the terms of the first group in the French corpus

In the French corpora (Fig. 6) the term “Enseignement à distance” (“distance education”) is used in articles, reports in 1970s, for instance, in the journal *L'éducation* (1971), volumes 99-105 digitized in 2011 (they belong to University of Pennsylvania).

Figures 1-6 demonstrate that in the Spanish, Italian and French corpora the phrase “distance education” was mentioned earlier than other terms; it is also the most frequently used one. In the other corpora the term “distance learning” dominates. Both terms are the most frequently mentioned in corpora in all six languages.

In the books corpora studied the concepts develop from general to more specific ones: from open and distance education to narrower and specific terms such as “distance courses”, “online courses”, which bears evidence of the tendency earlier revealed.

Without the most frequent terms of the first group dynamics of specific terminology development in education becomes more evident. Analysis of these terms statistics in the Russian corpus identifies the use of only two specific phrases (Table 1): distance courses and Internet education (Fig. 7).

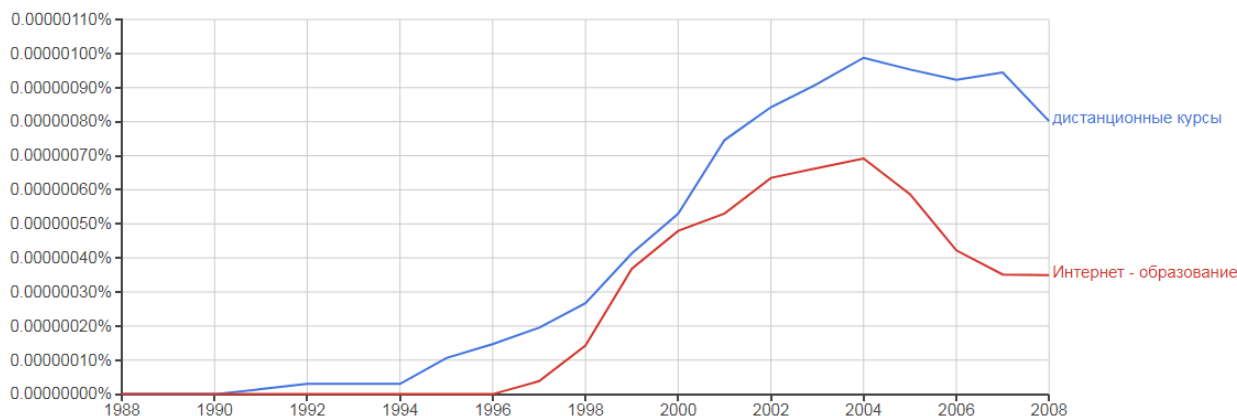


Figure 7. Frequency of the terms of the second group in the Russian corpus

The occurrence of these terms in the Russian corpus in early 1990s is caused by extensive use of computers and spread of the Internet. The terms are logically related to the concept of distance learning.

In the German corpus the phrases used vary more. The Ngram Viewer graphs the frequency of the terms “Fernkurse” (“distance courses”), “Online-Schulung” (“online training”),

“Online-Training” (“online training”), “Online-Kurse” (“online courses”), “Internetkurse” (“Internet courses”) (Fig. 8).

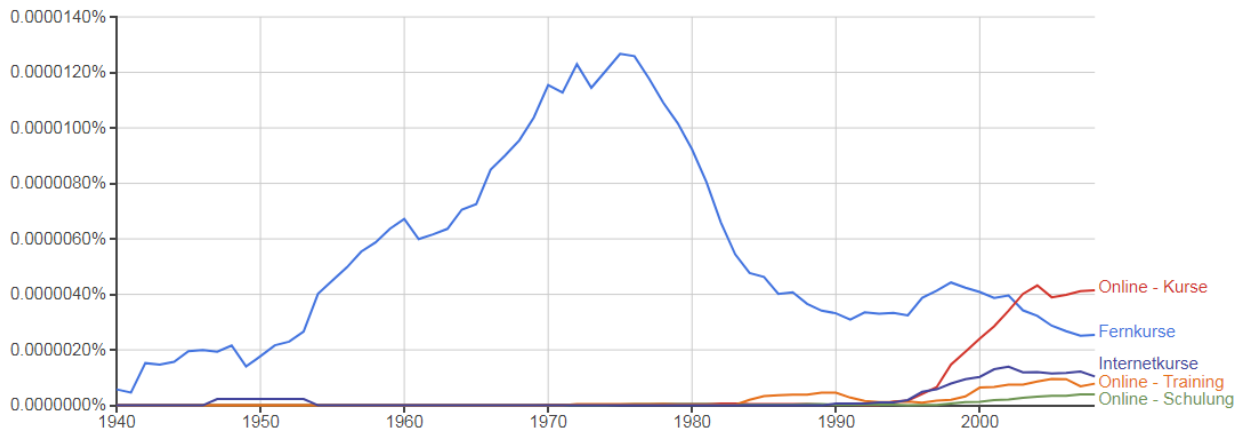


Figure 8. Frequency of the terms of the second group in the German corpus

Figure 8 shows that the phrase “Fernkurse” (“distance courses”) is predominant. The analysis of the sources where it occurs indicates that the first mentionings of the term relate to traditional extramural tuition. It is quite predictable that the use of the terms in today’s meaning starts in 2000s. It is the time of extensive use of computers and spread of the Internet. It is at this time that online courses become popular, which is due to the development of education in general and new ways

of getting education in different fields. It is notable that Internet courses do not only relate to University courses but to various workshops and training in a wide range of fields.

Similar dynamics of terms frequency arising from the progress of modern means of communication and technologies is seen in the other corpora. Thus, the English corpus reflects popularity of online courses and online education (Fig. 9).

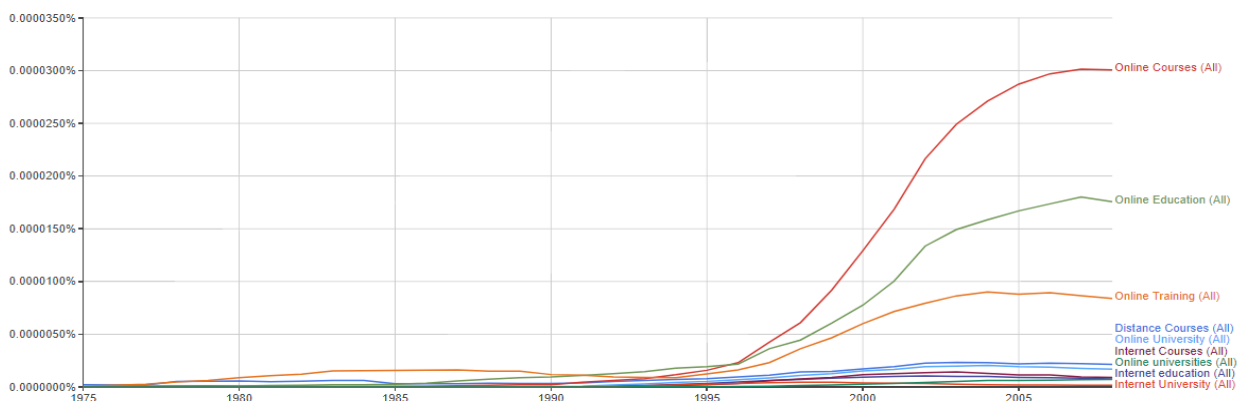


Figure 9. Frequency of the terms of the second group in the English corpus

The graph in Figure 9 shows classical logistic curves that hardly intersect. Smoother curves in this graph in comparison with the other graphs are indicative of the fact that the corpus of books containing the terminology studied is rather big, which makes it representative. Figure 9 clearly shows saturation points (approximately 2008) for the terms “online courses” and “online education”. It is the time when “The Theory and Practice of Online Learning” by Anderson [20], “Economics of Distance and Online Learning: Theory, Practice and Research” by Bramble & Panda [21] and other works were published. In fact, they link the concept of education to other areas of human activity, human and social development. The First International Conference “Hybrid Learning and Education” was held. It focused on hybrid learning presented as a combination of traditional face-to-face learning in classroom and Internet eLearning in the frameworks of one learning and teaching method, according to Fong, Kwan, & Wang [22]. Considerable attention is also paid to promotion of online education, for instance, the source DuVivier «100% Online Student Success» [23].

Graphs showing frequency of the second group terms in the Spanish and French corpora are similar to those showing their frequency in the German and English corpora. However, for some of the terms the Ngram Viewer shows only one source for each. The Italian corpus contains only terms “online courses” and “online education”.

Trend lines of all the graphs studied reveal that development of terminology is dependent on the needs of the society and the state, the educational policy pursued, social and economic environment for the development of educational system.

Along with traditional approaches more and more focus is directed towards developing education via distance Internet technologies. Distance education is becoming highly demanded as it provides universal access to new ways of getting knowledge.

#### IV. CONCLUSION

In this research the Google Books Ngram Viewer is used to analyse frequency of fifteen



terms representing the sphere of distance, open and online education and learning in corpora in six languages. For the purpose of receiving comprehensive results all the terms are divided into two groups: general terms and specific terms. The research shows that time period when the terms occur and develop is similar in corpora in different languages. However, there are works that the search in the Ngram Viewer turns up that are not from the sphere of learning and education. The terms used in such works are homonymous to those being studied. In the earliest works found, most of the phrases used represent the first group: general terms. The most frequent phrases in all six languages corpora are “distance education” and “distance learning”.

Dynamics of specific terms frequency reflects the progress of modern means of communication and technologies. Frequency of phrases with specific meaning peaks in early 2000s. Trend lines of all the graphs studied reveal that development of terminology is dependent on the needs of the society and the state, the educational policy pursued, social and economic environment for the development of educational system.

### REFERENCES

1. Keegan, D., 1996. Foundations of distance education. London: Routledge.
2. Bates, A. W. T., 1995. Technology, e-learning and distance education. London: Routledge.
3. Levy, S., 2003. Six factors to consider when planning online distance learning programs in higher education. Online journal of distance learning administration. 6 (1).
4. Baldassarre, M., 2016. Informazione, Conoscenza, Didattica. La sfida dei big data al mondo della formazione. Information, Knowledge, Didactics. The challenge of big data for the world of education: 90-112.
5. Silber-Varod, V., Eshet-Alkalai, Y., Geri, N., 2016. Analyzing the Discourse of Chais Conferences for the Study of Innovation and Learning Technologies via a Data-Driven Approach. Interdisciplinary Journal of e-Skills and Life Long Learning, 12.
6. Michel, J.-B. et al., 2010. Quantitative analysis of culture using millions of digitized books. Science, 331 (1): 176-182.
7. Jones, E., 2010. Google Books as a General Research Collection. Library Resources & Technical Services, 54 (2): 77-89.
8. Moskovkin, V. M., 2012. Google Books i “kulturologicheskiye trendy”. Scientific and Technical Information Processing. Seriya 1: Organizatsiya i Metodika Informatsionnoi Raboty, 7: 27-34.
9. Steklov, Yu., 1927. Mikhail Aleksandrovich Bakunin: yego zhizn i deyatelnost (1814-1876): Raskol v Internatsionale. Moskva: Izdatelstvo Kommunisticheskaya Akademiya.
10. Bolshaya Sovetskaya Entsiklopediya, 1934. Vol. 60. Moskva: Sovetskaya Entsiklopediya.
11. Vestnik Vysshoi Shkoly, 1947. Vol. 5. Moskva: Sovetskaya nauka.
12. Onushkin, V. G., & Ogaryov, E. I., 1995. Obrazovaniye vzroslykh: mezhdistsiplinarny slovar terminologii. Moskva: Rossiyskaya akademiya obrazovaniya.
13. Nyberg, D., 1975. The Philosophy of open education. London: Routledge & K. Paul.
14. Arrien, J. B., 1978. Los retos de la «educación a distancia»: I Seminario-Taller de Teleducación Universitaria de FUPAC, V Seminario Latinoamericano de Teleducación Universitaria. Federación de Universidades de América Central y Panamá (FUPAC).
15. Educación, 1974. Departamento de Asuntos Educativos, Secretaría General de la O.E.A.
16. Pintado, A. M., 1977. La reforma educativa de la Segunda República Española: primer bienio. Vol. 15. Madrid: Santillana.
17. Instituto Colombiano para el Fomento de la Educación Superior. Metodología y estrategias de la educación superior abierta ya distancia: nivel introductorio. Universidad Abierta ya Distancia, 1986.
18. Merlo, G., 1996. Formazione e lavoro femminile: il contributo del fondo sociale europeo: politiche ed esperienze. [Lavoro e società, 16]. Milano: Franco Angeli.
19. Quagliata, A. (Ed.), 2003. Pratiche di didattica costruttivista in aula e nella rete. Armando Editore.
20. Anderson, T. (Ed.), 2008. The theory and practice of online learning. Athabasca University Press.
21. Bramble, W. J., & Panda, S. (Ed.), 2008. Economics of distance and online learning: Theory, practice and research. London: Routledge.
22. Fong, J., Kwan, R., & Wang, F. L. (Ed.), 2008. Hybrid Learning and Education: First International Conference, ICHL 2008 Hong Kong, China, August 13-15, 2008 Proceedings. Springer-Verlag Berlin Heidelberg.
23. DuVivier, R. L., 2008. 100% Online Student Success. Cengage Learning.