THE LEARNING AND EDUCATIONAL POTENTIAL OF DIGITAL TOOLS IN HUMANITIES AND SOCIAL SCIENCE

Olga Suleimanova
Moscow City University, Russian Federation

Innara Guseinova
Moscow State Linguistic University, Russian Federation

Albina Vodyanitskaya
Moscow City University, Russian Federation

Abstract. The paper focuses on up-to-date cutting-edge digital technologies that may be used in teaching university students. The authors suggest distinguishing three types of digital tools, or engines: search tools, research tools and interactive ones. The didactic and research potential of these tools is analysed. The search potential of search engines and text corpora is compared, supported by practical templates which illustrate how the engines are employed to their best. The research potential is assessed on the practical example of semantic experimental research into the difference in the meanings of the words competence vs competency. The Google-supported experiment was elaborated and illustrated against traditional polls of native-speakers, the results of the two showed good concordance which testifies to the reliable validity of research engines for linguistic experiments. Special attention is paid to interactive engines which are used in the educational context, such as Mentimeter, or supporting an educational blog. The blog is treated as a powerful tool in promoting academic writing skills in students, their communication skills and a didactical tool used for sharing professionally relevant information among the students and professors. The blog elements are described.

Keywords: digital technologies, search engine, research engine.

Introduction

Digital technologies are omnipresent in our daily life, they permeate all human activities and are especially welcome in the youth’s lifestyle. Being a powerful source of communication and entertainment, they can serve as an equally powerful educational tool. The digital market offers a variety of search engines and research tools which the academics can choose from and recommend them to the student.

Application-wise, we will distinguish three types of instruments that can be used for teaching / learning purposes. These are: first, search tools – big-data bases (Google, Yandex and the like; national text corpora: BNC – British National
Corpus (Davis, 2008-), COCA – Corpus of Contemporary American English (Davis, 2008-), NCRL - National Corpus of the Russian Language (2003-), etc.); second, research tools, such as Vaal-mini (Dymshits & Shalak, 1992-), SentiStrength (Thelwall, Buckley, & Paltoglou, 2012), Tropes (1994), Mentimeter (Warström, 2014-) etc.; and third, the systems that can be treated as interactive ones. Their educational potential will be considered.

In this context the prime responsibility of professors is to research into abundant available web resources, assess them application-wise, i.e. understand the possible formats of applying them and potential problems that can be researched into using these resources. After that we offer recommendations to the students and instruct them as to how the tools can be employed. In some cases search engines will suffice, while in other ones research tools can be quite promising. What is new is the experimental potential of Google-type engines for linguistic research in addition to interviewing and sometimes substituting interviewing native speakers. Such procedures are waiting for their detailed elaboration and verification.

**Search tools**

Search engines are the most popular: we can rely on them in teaching languages (getting materials for the classes, exams, textbooks) and – less often – use them for research purposes: obtaining empirical data and processing them in relation to the “formal” characteristics (co-occurrence analysis, frequency-analysis), which can serve as a solid basis for content analysis, as a starting point for semantic and cognitive analysis (Rakhilina, 2017). Relatively new systems – compared to text corpora – used for linguistic search purposes such as Google, Yandex and the like, when used for such purposes, are arranged in the similar way and may offer new research perspectives compared to the corpora.

In dealing with these engines for linguistic purposes the researcher has to assess the correctness of the generated data and degree to which she/he can trust the obtained texts / statements, whether they are fully complying with the national language code, or not.

**Web or Corpus?**

The research potential of digital tools - still underestimated - for linguistic research, e.g. in cognitive studies for describing word meanings, has been emphasized in (Fischer, 2010, p. 44; Tummers, Heylen, & Geeraerts, 2005, p. 233).

Availability of two competing types of databases for linguistic research – text corpora vs web google-type engines – gave rise to debates which posed the
questions to follow: can web resources be treated as a full-fledge analogue of text corpora; what is the basic difference(s) between the two that can taken into consideration when choosing the best data base in solving research problems, and some other issues (Gatto, 2014).

Researchers claim that web resources offer wider scopes, variety and greater volumes of data, admit their high availability and accessibility, together with “impartiality” of the texts which are more free from the authors styles and choices compared to text corpora – the above features make up for the deficiencies that can be traced in the google-type web systems (Mordovin, 2015). However, as relatively new resources, search engines require further investigations, e.g. due to growing interest in them, lack of well-elaborated search procedures and detailed analysis of their full potential, assessment of the degree of trust in this data, their spontaneity and specificity.

One of the most widespread engines is Google – due to its high speed, simple operation, data capacity and easy functional information retrieval, although it is not free from some deficiencies such as high information noise (which can be easily removed using mathematical probability theory – see (Fisz, 1963)), loss of some information, etc.

I.M. Petrova offers a comprehensive comparative analysis of the web vs text corpora research potential with reference to binomial noun phrases (Petrova, 2018). She analyzed the Google logistics and described the role of the search thesaurus (special dictionaries for information retrieval, a descriptor as a means of lexical control – e.g. detecting homonymy, synonymy), an internet flexible search algorithm which relies on operators, symbols. This means help modifying the search criteria, as well as their combinations, which allows to enter complex queries. For example, operator “… ” fixes the order of words in a phrase, or the keyword encoded as a direct request, or as incorrect entry. In the latter, the phrases are changeable – they admit other words, varying word forms, word order and punctuation marks. The more detailed the retrieval request, the better the resulting data.

I.M. Petrova concludes that web resources offer practically unlimited data, belonging to a great variety of stylistic genres; when preparing information for the web, it is processed by the system in terms of semantics, a wider search function allows the user to get more focused data to meet the more demanding needs. The text corpora are less flexible, the data are limited in some respects (Petrova, 2018).

It is clear that web resources are a promising tool calling for further development as a search engine.

Both data bases can be used as research tools in experimental / empirical research.
When employed as a research tool, Google, Yandex and others are actually used in experimental format, for semantic / cognitive purposes (Suleimanova & Petrova, 2018). It means that they can substitute native speakers of the language who are interviewed in the experiment and work as “interviewees”. Let us demonstrate how it works using a pair of words competence / competency and exemplify how databases can be used for semantic analysis (full comprehensive semantic analysis of these words can make the subject of a special research and is not offered within this paper). These words are actually treated as absolute synonyms, they are defined in the Longman dictionary: competence also competency 1. ability to do what is needed (the second meaning refers exclusively to court and is a legal term, the third is qualified as lit or old use – these two meanings are not analyzed here). In VisualThesaurus the words are also explained one through the other. What follows from the definitions is that the two words share the same meaning and there is no difference between the two, which can be challenged as linguists agree that there are no absolute synonyms in the language. If there were, we could expect that the words could be used interchangeably in all the contexts. Working on that presumption, we tried to substitute one for another in some contexts and discovered that they were not completely interchangeable, e.g. we can say I admire his competence as a supplier and cannot say*I admire his competency as a supplier. (Here the informant said that Being a supplier is a generic activity, and cannot be described as a specific skill.) So we conclude that there is some difference, which has to be researched into.

Here we start working with the text corpora – first in quite a traditional, well-elaborated format, calculating their right and left distributions, as “the complete meaning of a word is always contextual, and no study of meaning apart from a complete context can be taken seriously” (Firth, 1935, p. 37). Moreover, the ‘bottom up’ method (J. Channel) is efficient when “researchers must have in front of them a large number of examples… the researcher cannot start by ‘thinking of an example’ and then look for citations of it. Hence observations … involved looking in detail – and individually – at most of the lexis of current English” (Channel, 2003, p. 41).

Corpus experiment

The research format began with searching BNC (British National Corpus) and COCA (Corpus of Contemporary American English) for utterances which contain words competence and competency and calculate their distribution characteristics.
The frequency analysis showed for competence in BNC 1482 occurrences; in COCA – 7053 occurrences; competency – BNC only 58 occurrences and in COCA – 2110 cases. One might presume that, first, competence is more widespread in the American English (Am.E) than in the British English (Br.E); second – competency is for some reason (which we hope to find) is quite successfully competing with competence. Still, competency is far less frequent (esp. in Br.E).

Corpora data can be substantially complemented by web data, i.e. Google Books Ngram viewer (Davis, 2011) shows that competency used to be more frequent since 1800 than competence. First digitally recorded mention of competency date back to 1800 (A Report of the Debate in the House of Commons of Ireland) and it was mainly used in court. It sheds some light on the difference between the two words – competency tends to refer to professional discourse, e.g. in the legal language. It used to denote financial background of a person and gradually started since the 19th century to develop a more general meaning. The formerly infrequent competence rapidly grew more frequent.

See the graphs below.

*Figure 1 The dynamics of the use of the word “competency” years 1800-2000 (Davis, 2011)*

It shows that competency does not demonstrate significant fluctuations and remains within the statistical error.
The graphs show the relative stability of competency as compared to competence which is gradually becoming more “active” since 1940.

In this way we found one more argument which supports our claim that the two words are not identical in their meaning.

Proceed to the next step -corpora analysis and calculate distributions of the two nouns

We found in COCA and BNC that competence can be preceded by a variety of adjectives, such as special, professional, linguistic, social, basic, potential, communicative, and some noun attributes: vision / language / teacher competence; it is governed by such verbs as limit, display, strengthen, lack, achieve (the) competence; competence may be in some sphere: in English / using environment / in moving / in personal mobility.

Competency combines with the attributes teacher competency (test), in stock phrases competency based training methods / curriculum / testing; can be governed by verbs develop, enhance, assess, achieve, etc.

What follows is that both words share some distribution features: they both can be modified by the adjectives cultural competence / competency, teacher competence / competency, be governed by the same verbs. And now we can exploit Google as a research engine and “ask” it if it is possible to use them interchangeably in all the contexts. We selected minimal word combinations, substituted the original word with its synonym and sent a request into Google which provided, first, statistics and a group of combinations which Google rejected and suggested using competence instead of competency: potential competence of (3 510 entries) while for potential competency of the system suggested using potential competence. See also recommended change of degree of competency for degree of competence, or language competency for language.
What follows from statistics is that competence nowadays is much more frequent, besides we were able to generate a hypothesis (which was later supported by an interview with a native speaker (Robin C., aged 73, degrees in Modern History and Psychology, Market Researcher and former EFL teacher, now retired) which runs as follows: judging by the usage and Google-based focused experiment where the speaker discussed the process of education, which can be measured, tested (it is naturally based on step-by-step mastering of different skills), the team competency matrix, competency based training (720 000 entries, in contrast to competence based training – decimated to 74 600 entries), competency levels 303 000 (cf. competence levels 153 000), or competency chart 1 410 (cf. competence chart 415), competency based 6 270 000 (against competence based 1 030 000), competency test 559 000 vs competence test 104 000, computer competency 168 000 and computer competence 63 700 vs competency in 6 340 000, competency model 546 000 vs competence model 214 000, we presumed that competency refers to some skill which the protagonist can obtain and employ, while competence refers mostly to the general ability to act as required in the given context.

Not surprisingly, our interviewee (we referred to several informants and asked them to assess the correctness of the phrases in accordance with the well-elaborated procedures of the semantic experiment (Seliverstova & Suleimanova 1985), using 3 grade scale. Some of them assessed the phrases, others, like the abovementioned one, tried to offer their ideas which were analyzed and related to our hypothesis – see below) claims that the word “competence” means having the ability to do something at an acceptable level. By extension it can have a legal meaning that it is an area over which a person or body, institution etc. can have the power or authority to act. He suggests an utterance where both words are interchangeable “We would expect a person coming to our house to service the central heating boiler to have competence in plumbing / “The company expects employees servicing central heating boilers to have competency in plumbing.”

The next explanation the informant suggested is different: “Many people dispute whether the decision to prevent Boris Johnson from proroguing Parliament lies within the competence of the Supreme Court, because it was a matter that should have been decided within Parliament and should not, in theory, be a matter over which the Law has jurisdiction.” The example above about Parliament is logically not possible if we use “competency” instead of “competence”. Competency is limited to a specific skill or ability. So I don’t think “competency” would make any sense in the Parliament example.

Comparing the data we obtained from the database and the hypothesis we suggested, on the one hand, with the poll of the native speaker(s), on the other,
we see that healthy combination of traditional semantic experiment interviewing native speakers and research engines empower the researcher with a powerful time-consuming triangulation instrument if used in the experimental format (Suleimanova & Demchenko, 2018).

Then, other research engines, to mention such programs as, e.g., Vaal-mini (determines the emotional charge of the text, works on Russian texts only), SentiStrength, Tropes, Google Scholar, Trello, Mendeley, Scrivener, Scanmaker empower the researcher with the mechanisms which are focused on specially designated tasks, e.g., determine the potential emotional impact on some target group, carry out linguistic expertise of the text (in practical purposes – in court, in analyzing potential terrorist threats, etc.). These tools are actively used by the linguists as they show quite reliable results. (We recommend these tools to the students and they, as “digitally born” millennials (Sheninger, 2014), and many of them already computer geeks, refer to them and successfully employ in their research. These programs are free and available in the internet, are easy to employ and are accompanied by clear instructions. One has to enter the program name and will immediately get an access. What the professor is expected to do here is give a student some guidelines and the program name which will direct the student towards the relevant tool.

Of special interest is Mentimeter as a psycholinguist’s helping hand / as an instant poll results provider. It is a relatively new research tool which generates polls’ returns immediately in 12 different graphic formats, e.g.

![Mentimeter poll results presented in the form of a word cloud](image)

*Figure 3 Mentimeter poll results presented in the form of a word cloud*
This word cloud was generated in the poll at the 54th Linguistics Colloquium (Moscow City University, September 2019). The picture with the results of the poll features the information about the number of participants (124) as well as the platform (Mentimeter). (After the poll is completed the results are sent by the system to the e-mail address of the author of Mentimeter presentation.)

And that makes it an indispensable tool for psycholinguistics where a researcher, instead of asking (often reluctant) respondents to kindly participate in a poll, can get immediate returns from willing and involved participants who actually enjoy, first, being polled via the phone (inclass), second, watching the result immediately dynamically featuring in a very attractive format on the screen, in the form of a constantly changing cloud. The process is started with the command to *take and switch on the phone*, then go to www.menti.com, then enter the code 11 22 57 and vote. Voting can be carried out in the digital or verbal form, e.g. a researcher— a psycholinguist, or a social scientist, a specialist in (cross) linguocultural studies—is studying associations with some concept that are generated in the consciousness of the interviewed. The latter is asked to punch in the word-associate(s) into the phone. As easy as that! The results are being displayed on the screen online with the polling process going on in the real time. Participants to the poll (usually young) are fascinated by the process, being immersed into it. Students admit that they enjoy this process both as researchers and as the interviewed. Third, this program offers 12 (!) possible graphic representations which are automatically generated by the system, to add to the variety of the learning format. This tool’s potential needs to be practically analyzed yet.

**Interactive Engine**

The third of the variety we are promoting here is even more recent (see Sheninger, 2014) and we would like to open a discussion on how we as teachers can benefit from it and demonstrate it on the practice introduced in Moscow City university.

It is running blogs, we practiced two independent ones – one for the students (actually monitored by a student) and one more for the teachers, with the educational message (see https://edublogs.org/?join-invite-code=7190279-65). We argue that such blogs differ from those run on one’s personal experience, on what the bloggers saw and did. Educational blogs are meant to promote educational purposes and to disseminate professional knowledge; the blogs help students master relevant skills which they acquire both while actively contributing to the blog content and to downloading the experience and expertise from the previous learners.
The student and professor blogs are mirroring each other to some extent as they share some materials. They comprise the following parts: Academic Chances, Academic Faces, Academic Tips, Academic Sharing, Academic Events, Academic Classroom English and some others.

What we had in mind when we started the project was the idea to boost students’ and teachers’ creativity and motivation engaging them in contributing to the contents, in addition to providing them all with the relevant information concerning learning.

For instance, 2nd year students were asked (within the General Linguistics course where they were discussing research methods and techniques, in view of the term research paper) to reflect upon the digital methods they will use in their research, in the form of an essay (500 words). The essays were sent to the blog, and the students will have a chance to read their peers’ essays, moreover, next year students will be in a better position as regards their progress as they will be able to relate the information to their own research without teachers’ instructions. It means that the students are actually teaching each other, and learning from each other. In fact, first-year students are already using this blog element (in Moscow City University students majoring in translation studies are assigned research problems in their first year which are to be presented in the form of a term paper). The students share their experience in the form of essays – e.g., undergraduates write about their educational trajectory, about what motivates them in their studies and career aspirations. These essays are in great demand with the first-years and sophomores who are developing their professional identity through them.

The point is that the blog implies regular feedback and the practice of contributing there is gaining momentum – students find facts and figures which are relevant for their professional growth, and this practice of sharing knowledge boosts their professional self-esteem. It makes students more sensitive and responsive to the information around. Besides, developing academic writing skills is professionally relevant for prospective translators.

The blog can be also helpful in many other respects: e.g. Academic chances features fresh information on forthcoming academic events (conferences, publications), both for teachers and students, for example, the most recent information was on upcoming in April, 2020 “International Science and Practice Related Conference “Relevant Issues of Variantology, Communication and Cognitive Linguistics”. Despite the call for papers, the blog features the registration form for this conference and the form to fill in submitting author’s details, which is very convenient. Academic faces provides information about teachers’ publications – it can be used both by students, colleagues, and third parties who decide on the chair as a collective opponent for the thesis, etc. When students embark on the research career they have a happy chance to get a comprehensive picture of the faculty / chair members’ research interests, of what
is being done by their professors and immediately feel involved and “at home”. Besides, they feel free to participate, to contribute to the common cause. For example, this category features tips and guidelines on writing Web of Science- and Scopus-indexed papers. *Academic Tips* deal with (random) new and promising research vistas and cutting-edge research (web) tools, description of the British National Corpus, COCA, Google Ngram Books, tips on Mentimeter usage and guidelines on empiric base description. *Academic sharing* offers synopses of the professors’ achievements which briefs on three key innovative results of their papers. One can find there information on Web of Science publications, papers, monographs and other relevant publications. *Web Academic events* focuses on the educationally-relevant forthcoming events in the local and global educational environments, meant for both the professors and students. This year Moscow City University celebrates its 25th anniversary, so as a part of celebrations and as a part of International Translator’s Day, students of MCU presented their posters focusing on various types and aspects of University activities (international cooperation being among them). *Academic classroom English* is devoted to phrases, words and communication strategies which lead to the best perception. For instance, instead of saying *I am not ready* the students are encouraged to say *I’d rather wait before I commit myself*, which might as well be relevant when they are qualified and move up their career ladders. The question *Who is missing today?* at the beginning of the lesson sounds tolerant and up to the time. The best answer that might follow is at least *P. is missing*, with a still better phrase *P. sends his / her apologies* which might save colleagues’ reputation when used at work and at the same time produce a favourable impression on the potential stockholder.

*Academic malpractices* treats translation mistakes which both professors and students observe around, it helps boost professional sensitivity to the wrong practices and finally avoid and prevent them. Moreover, here researchers can find relevant guidelines on avoiding common mistakes in academic writing (*следует отметить* instead of *стоит отметить*).

**Conclusion**

In the digital era, against the background of abundant web resources, we have to choose the ones which exactly meet the requirements we set and we must be able to assess their potential and efficiency for solving the problems we are facing in the educational practice. We argue that, application-wise, three types of instruments that can be used for teaching / learning purposes must be distinguished. These are **search** tools – big-data bases (Google, Yandex and the like; national text corpora: BNC, COCA, National Corpus of the Russian Language, etc.); **research** tools, such as Vaal-mini, SentiStrength, Tropes,
The search tools represented by big-data engines such as text corpora and databases of the Google type are used to service traditional purposes – they provide unlimited text information, more abundant and well classified in Google as a rapidly developing engine and similar data bases as well as text corpora.

Using research tools still needs elaboration and assessing their potential for (student) research, especially for experimental linguistic research, where it is necessary to combine this potential with traditional polls of the native speakers.

Interactive tools are powerful motivators for the students and contribute to a variety of professionally relevant skills, such as academic writing, research skills, professional communication formats, when students get involved into profession, its challenges and prospects as early as possible through communicating with the peers and professors.

Distinguishing three types of digital tools, or engines: search tools, research tools and interactive ones and analyzing their didactic and research potential, supported by practical illustration of how the engines are employed to their best let the authors demonstrate that digital technologies being a powerful communication and entertainment resource in addition may serve as an equally powerful education tool. The (student) researcher can try and combine the tools, taking into account their deficiencies and advantages, understanding their pros and cons. The digital market offers a variety of search and research engines which the academics are to choose from and to recommend them to the student.

Acknowledgements

We owe a debt of gratitude to Robin Carr who agreed to take part in the interview as a native speaker of English and suggested comprehensive in-depth commentaries that nicely comply with big data results we have obtained.

Literature


