English-language Digital Discourse of Human-Machine Communication

Nataliia Lazebna,[*] Anatoliy Prykhodko[†]

Abstract

The paper focuses on digital discourse. This is a speech-intellectual product of innovative information technologies, a phenomenon, which needs further interdisciplinary and linguistic interpretation. The English-language digital discourse shows how linguistic verbal communication is mediated by digits and to what extent these Signum and Verbum unity reigns over the world.

The paper analyzes the ways and methods of integrated and differential use of verbal and non-verbal sign systems in the English language as compared to programming languages, considering the types of synchronous changes in the socio-cultural dimension of the sign. This research describes the processes of signs transformation during their functioning in programming languages and in the English language, common and distinctive features in the arrangement of grammatical, lexical-semantic, and graphic means of (natural) English and (artificial) programming languages in their projection on different modes of communication in the system Human ↔ Machine.

Programming languages are constituted by verbal means of the English language with additional use of its own semiotic resources, which testifies to their integrative linguistic and mathematical nature. The specific representation of ElDD conveys its reciprocal nature when the English language using its own tools combines them with the elements of the programming languages thus creating an effective toolkit for self-processing.

Keywords: English-language digital discourse, linguistics, semiotics, digit, text processing, programming languages.

The scientific and technological progress of modernity disturbs the monopoly of natural language related to human communication, developing and contrasting it with programming languages. The linguistic dualism «natural language – artificial/synthetic language» is today an inherent feature of the information society, which begins to form the «machine picture of the world» (Jokuza, 2017, EP). English-language digital discourse (ElDD) outlines a new status of a Human in the form of «intermediate existence» between the real and virtual worlds. The phenomena of «naturalness» and «artificiality» are interrelated and in the

---

[*] Privat Dozent, Dr. Habil., TEFL Methodology Department, Julius-Maximilians University of Wuerzburg, Wuerzburg, Germany.

[†] Full Professor, Dr. Habil., Chair of Translation Theory and Practice Department, Zaporizhzhia National Polytechnic University, Zaporizhzhia, Ukraine.

context of studied discourse are harmoniously intertwined. The above mentioned pre-requisites of the research create an innovative basis for new linguistic development, digital linguistics, with ElDD as the central concept. ElDD is represented as a three-layer cognitive and communicative construct in the unity of micro-, meso-, and macro levels used for Human-Machine communication.

The Empirical Material is represented by English-language texts and text fragments, which relate to the existence of ElDD. The first part comprises texts in the English language, which are used in programming languages, for example, Python guidelines with examples [2012-2020], Scratch guidelines [2008-2010; 2020], and Squeak by Example 5.3 [2010-2020]. The second part includes texts used in innovative technologies focusing on texts describing natural language processing Multi-Domain Wizard-of-Oz dataset, NPS Chat Corpus, Stanford Question Answering Dataset (SQUAD). The third part comprises texts, which reflect the use of technologies for language processing (Github.com – scripts and answers of chatbots Alice, Eliza, Mitsuku, Rose).

Methods:

- interdisciplinary method is implemented to consider ElDD in different fields of science;
- descriptive methods to interpret and generalize basic practices of modern discursive studies;
- semiotic analysis to parametrize linguistic relevant properties of programming languages and natural language;
- linguo-cognitive analysis focuses on communication triggers between Human and Machine;
- discursive-textual analysis focuses on ElDD specifics and different levels of textual representation;
- functional-communicative method to analyze grammatical, textual, and graphic properties;
- structural modeling is used to grammatically describe the processes for clusterization, templates, and tokenization of ElDD textual space.

General Discussion

Pathos and mystery of human communication are broken by a third party, a digit, which becomes a kind of «superstructure». The natural language is specific and flexible because it quickly adapts to the laws of information technology and thus appears both as a primary system, a donor of discourse, and as a secondary modeling system or a focal point for creating artificial languages. ElDD is embodied as a powerful linguistic, social and cultural phenomenon (Lazebna, Prykhodko, 2021, EP).

ElDD, at the same time, is a kind of institutional communication mediated by a symbiosis of (artificial) programming languages and (natural) English language, which reflects

the digitized speech act motivating verbal interaction of its agents and its clients by means of innovative information technologies. This studied type of discourse is an ideal self-generating system. Its linguistic and semiotic dominants focus on coding and decoding commands both aversively (from Human to Machine) and reversibly (Machine to Human). In this way, there is a reciprocal interaction between the target sphere of discourse (natural English) and the source-sphere (programming languages).

String»lineCount
"Answer the number of lines represented by the receiver, where every cr adds one line".
/cr count/
Cr :=Character cr.
count:= 1 min: self size.
self do:
[:c / c == cr ifTrue: [count := +1]] (Squeak, 2021, EP).

Fig. 1. Python language of programming

Figure 1 illustrates phonetic, lexical, and grammatical relations, which are transformed into a kind of linguistic and mathematical symbiosis, which is successfully used by discourse participants, its agents, and clients.

Participants of ElDD communicate within the appropriate text space, filled with digitized texts of clearly structured composition, certain cognitive-semantic reflections, and specific pragmatic guidelines that require appropriate competence. Textual representations of ElDD are a kind of intellectual products of participants' mental and speech activities in at least three registers of communication («Human–Human», «Human–Machine», «Human–Machine– Human – Human») and their variations, depending on the modes of direct or indirect communication, as well as the mode of their encoding and decoding.

For example, an abstract from the programming language Squeak highlights two modes of interaction «Human User – Human Programmer – Machine – Human User», where the User is looking for an explanation of the programming language functioning:

1) **FAQ 5** The browser does not look like the one described in the book. What gives? **Answer** You are probably using an image in which the OmniBrowser is installed as the default System browser. You can change the default by clicking on the little menu icon in the top-left corner of the browser window, between the «X» and the browser label, and selecting «Choose new default browser» from the menu that appears. Select #Browser to get the plain old browser. (This preference changes the effect of the World ◁ open . . . ◁ system browser menu, but not what happens when you drag the Browser icon from the Tools flap.) (RecipeQA, 2019, EP).

2) **FAQ 10** How do I tell which methods of a class are abstract? Answer abstract Methods := [:a Class | a Class method Dict keys select: [:a Method | (a Class>>a Method) is Abstract]]. Abstract Methods value: Collection – → an Identity Set (#remove: if Absent: #add: #do:).

In the first example (1) the author of the guideline explains the functionality of the programming language Squeak, using means of the natural language and only rarely built-in

commands, symbols, and signs. In the second abstract the author appeals for the commands of the programming language Squeak.

The interaction «Human↔Machine» will be effective when the sender achieves his goal and the recipient meets the needs of the sender, but it does not always work, because the effectiveness of this interaction depends largely on the Machine. By changing voice commands, or making some phonetic or typographical errors in a search engine (such as Google), the Machine either modifies or cannot process the sent command / request, because the coded nature of the Machine depends on the dynamic and changing nature of Human (Makhachashvili, Semenist, Zatsnyi, Klymenko, 2021, p. 148). That is why, it is impossible to modify and improve the acts of communication between Human and Machine, focusing exclusively on linguistic and mathematical procedures. Here, it will be necessary to re-deploy the interaction «Human –Machine»and turn it into a relationship «Client ↔ Programmer ↔ Machine ↔ Human»/ «Human-1 ↔ Human-2 ↔ Machine ↔ Human-1» / «Human-1 ↔ Human in Machine».

Whatever register of discursive communication between Human and Machine is considered, it will always produce the text as an output. The text space of English-language digital discourse has a three-layer structure, which consists of macro-, meso- and micro-levels.

Structure of horizontal contexts of ElDD

Texts of macro level			
Layer 1	*Layer 2*	*Layer 3*	
Verbal reflection of the information register	Meaning	Image	Horizontal context, commentary, hyperlink, navigation buttons
Texts of meso level			
Verbal/nonverbal reflection of the information register	Meaning	Code Tag	Vertical context, hyperlink
Texts of micro level			
Verbal reflection of the information register	Meaning	Image	Vertical context

Table 1. Structure of horizontal contexts of ElDD

The *microlevel of ElDD* is based on the texts of programming languages, lists of commands, and written programs having a mathematical basis with its set of digits and symbols. On the one hand, programming languages have the status of artificial ones, as they are chains of predications and written commands, and, on the other hand, they are a powerful tool for constructing digital discourse. In interaction with natural language, they form the inner linguistic- semiotic contour of English-language digital discourse.

The emergence of programming languages as a kind of «unified digitality» based on natural languages was a radical breakthrough in the movement of Human to artificial intelligence and opened new opportunities for communication with the Machine. On the one hand, the English language appears in this system as the primary donor system of programming languages, in particular, and digital discourse in general, and on the other hand, represents the secondary modeling systems that create artificial languages. In this context, the English language develops human-coded world and its fragments (Black et al., 2020, p. 223).

The most common programming languages based on the English language are Squeak, Scratch and Python. They create the internal contour of ElDD and have no oral channel for transmitting information, these languages are interconnected by global compositional determinism at the algorithmic level. Their differential feature is the lack of a common domain of interpretation between programs written in the natural (English) language. Most of them are characterized by local variables, systematic indexing, recursiveness, and lack of semantics without syntax.

Linguo-semiotic elements of micro level produce networks of relationships in which each of these elements in the context of programming languages or natural (English) language can be regenerated from a word or symbol into a command or query. For example:

```
# Define the sentence to be lemmatized
sentence = "The striped bats are hanging on their feet for best"
# Tokenize: Split the sentence into words
word_list = nltk.word_tokenize(sentence)
print(word_list)
#> ['The', 'striped', 'bats', 'are', 'hanging', 'on', 'their', 'feet', 'for', 'best']
# Lemmatize list of words and join
lemmatized_output = ' '.join([lemmatizer.lemmatize(w) for w in word_list])
print(lemmatized_output) (Scratch, 2021, EP).
```

In programming languages, communication is mediated by both machine and software codes. Therefore, visually (in the sign-symbolic version) you can see only a part of the information technology, i.e. the external form and user interface. In this case, the system of open-source code, algorithms, and structured representation of databases, technologies, techniques, and interfaces enables effective communication between Human and Machine (Gee, 2014, p. 134-147).

Texts created on the basis of programming languages show the clarity of their natural language form, differ in the nature of the prescribed commands and operational order. Further, these texts penetrate into the fabric of natural language and are embodied in the digital discursive environment mainly focusing on the use of a certain sequence of elements. These include lists, lines, tuples, and elements of the dictionary. They are all grouped around the idea of variability, the immutability of data, and orderliness/disorder of elements. In general, the textual representation of the micro-level of ElDD is characterized by structure, coherence, cohesion, the logic of construction and representation, as well as a symbiosis of natural language and (sub) language of mathematics (Gabbrielli, 2010, ER; Herring, 2004, p. 316).

The meso level of ElDD is the center of direct interaction of natural English and artificial programming languages. At this level, the former, using its own tools, integrates them with the elements of the latter, creating effective tools for processing «itself». Initially, the natural English language builds up the «bricks» for the development of analytical tools (creation of chatbots, natural language processing programs, NLTK, etc.), and as a result, the English language itself becomes the object of processing. In this way, the natural English language is regenerated in a technocratic and digitalized space.

The leading morphological categories most involved in the interaction of English and programming languages are auto-semantic units, which are, first of all, noun and verb. A nominative phrase with a noun at its top has a wide functional range for the transfer of information from the Machine to Human and vice versa. The nominative focus of ElDD participants initiates the activities of Human in cyberspace (creating requests for the Machine, searching for information in the network, verbalization of commands of programming languages, etc.). In combination with the verb, which is the vehicle for the categorical meaning, the effectiveness of information transfer increases many times. The verb appears in this system as a key representative of the message template, mediating the so-called agent focus. The agent focus of programming languages «migrates» from the ascertaining to the restrictive, not limited to the given semiotic framework (symbols, marks, tags). Within agent focus, commands of programming languages can be split into different constructs of alteration, creation, establishment, and obtaining. For example: *create_ package, generate_statement, start_listening* etc.

Morphological «awareness» of Human and Machine plays an important role at the meso- level of the analyzed discourse. It conveys the mastery of the mechanisms, which break tokens into morphs, which in the process of communication helps to «distill» the information transferred to the machine (Farhadi, 2015, p. 187). Morpho split is a process and, at the same time, the result of the «distillation» of morphs in the process of communication. Due to its digital nature, the morpho split becomes an important prerequisite for text clustering. The machine proceeds to the clustering of textual matter, triggering the morpho split. At the same time, it relies on a huge sample, turns to previous developments and vector representation of words, and involves tokenization, purification, and stemming to remove «noise» signs and symbols. For example, there are possibilities of active request representation (*classification, elaboration, research, treatment*):

Human: *Can you block my debit card?* Machine: *You're all set, your card is now locked. You can unlock it anytime with my help or in the app*.

Another option is to **represent characteristics of the request subject** (*attitude, consideration, discussion, reason, sense, availability, complexity, feature, order, reaction*): Machine: *Hello. How may I help you?* Human: *Hey, I lost my access.*

In ElDD there is also a description of request (*brief description, explanation, presentation*). Machine: *Here's the sweatshirt. Click below if you want more info, if you want more info or buy it.*

For the Machine as an agent of the meso-level development, grammatical categories of natural (English) language are transmitted using digital identifiers (gender, digit, time, etc.) Two important aspects of the interaction between natural and programming languages are temporality and modality (Van Leeuwen, 2011, p. 688).

In the case of text processing by means of the English language, the machine clearly follows the established algorithm: text → random separation (cleaning) of lexical units → ordering lexical units by parts of speech → digitization of lexical units (digitization). In the process of natural language processing, which is both an object and a tool of analysis, there is a need to algorithmize the mechanisms of machine word processing. Purification strategies become very effective: representation of the request, description of the request, means of information transfer, etc. In this process, the role of the Machine becomes exclusively instrumental: it performs its analytical function according to the criterion of words used. Ambivalence, synonymy, homonymy, and polysemy become interfering factors in the process of natural language processing by the Machine.

At the macro-level (external linguistic and semiotic contour) of its representation, ElDD creates the environment, mode, and style of sharing natural English by Human and Machine to establish mutually beneficial communication by producing structurally integrated, meaningfully relevant, and pragmatically targeted speech messages (Lazebna, 2021, p. 175). The mediator of such communication is a chatbot, an artificially created program that simulates interactions/conversation/dialogue with the User. The chatbot is gradually becoming a universal mediator in Human↔Machine communication. A chatbot is a certain «anthropomorphized» entity, created by a programmer, and is not just some digitalized entity, but something «humanlike». Chatbots range from animated interface agents, conversational agent and avatar. In terms of ElDD, the Machine can be both an author and interlocutor. A User interacts with the Machine to meet personal needs, avoid feelings of loneliness, search for information, and so on. The effectiveness of chatbots mediated by means of natural language is supported by introductory framing, special triggers, templates, and techniques for eliminating syntactic and semantic ambivalence. Chatbots involve a certain tactical and strategic arsenal, such as return, clarification, paraphrasing, reorientation, and so on (Yan, Duan, Bao, 2016, EP).

When the User communicates with an animated conversational agent, the model of living / natural speech is reproduced. Its successful pre-requisite is the operational linguistics of discourse, the logic of using appropriate elements of programming language, the efficiency of the Machine response, and most importantly, the response of the «Machine» to «Human». Similar to natural communication, such speech exchange involves a rhythmic alteration of the communicative roles of the sender and the recipient, the client and the agent of the discourse.

The communicative and functional potential of discourse plays an important role in the management of human and machine speech exchange, which involves the harmonious use of compositional and architectonic-speech forms, speech and discursive acts, maxims of cooperation, and politeness. With the proximity of natural («human») communication, the macro-level of ElDD actively uses established architectonic-speech forms (monologue, dialogue, polylogue) and the classical «quartet» of illocutionary acts (declarations, directives, commissives). In terms of this harmonious unity, discursive acts (offeratives, con-formatives, accusatives, apologetives, didactives) are developed. The most frequent are offeratives.

Within these speech acts, the Machine implements the offer of services provided to Human. Among the discursive acts that make up the text space of ElDD, prohibitives are not

frequent ones because their use is associated with the categorical nature of the Machine, with its frequent «inability» to meet the needs of human.

The act-speech pattern of communication between Human and Machine at the macro level of ElDD is created in at least three ways. First, it is a strategy of effective communication (return, clarification of the request, paraphrasing, simplification), and secondly, it is following the maxims of cooperation (quantity, quality, relevance, method), and thirdly, compliance with the principles of courtesy (tact, approval, modesty, consent). These maxims establish effective communication. Humans rely on their own cognitive-sensory experience and the machine relies on prescribed algorithms, contexts, and templates. In this communication, there are at least two contours, both internal and external. The first is the contextual predetermination of communication, the second is the contour of attraction, which is reflected by the Machine in acts of friendship, sociability, and ethics. In recent years, a special code of ethics for the Machine has been developed, which internalizes impartiality, correctness, reliability of the information, and respect for Human. It also covers issues of privacy and humanity (objectivity of information, adherence to political correctness, avoidance of gender, racial, age discrimination, etc.). Following the principles of Ethics of the Machine, there are certain issues covered:

> Avoid stereotypes *What's your gender?*
> Interpersonal relationships *I am a human too.*
> Personalization *I am a Human. I am not a robot.*
> Peacemaking *Keep calm! Relax. Everything will be fine*
> Humor and friendly communication *Do horses go to Harvard? Horses go to Hayvard*
> (Cleverbot, 2021, EP).

Multimodal interaction of Human and Machine contributes to the gradual formation of a digitalized global world, where the boundaries between Man and Machine are blurred. Animated agents are no longer some images on a computer monitor, but are the embodiment of some anthropomorphic entity, a virtual but equal interlocutor of Human. Going from a mechanical informant and/or automated assistant to a full-fledged interlocutor, the chatbot is gradually transforming into a socially active and sound individual.

Conclusion

The proposed cognitive-communicative analysis of English-language digital discourse may in the future become the basis of linguo-synergetic, linguo-conceptual, linguo-cultural, and linguistic-stylistic descriptions of other digital discourses on the material of different natural languages. The implemented analysis would be helpful for analysis of other discursive formations, conceptual pictures of the world, cognitive and functional comprehension of text spaces, and contrastive research reflections, including translation studies. The phenomenon of ElDD can and should become a source of in-depth cognitive reflections in different fields of humanities, such as psychology, philosophy, cultural studies, and computer science. The introduced approach can also contribute to the expansion of knowledge and ideas of the linguistic community about the role and place of ElDD in neuro-linguistic programming and language manipulation techniques.

References

1.   BytePython. URL: https://pythobyte.com/python-keywords-identifiers-7cd08b65/ (accessed: 27.01.2021).

2.   Cleverbot.   URL: https://www.cleverbot.com/j2convbydate-qg14955   (accessed: 15.05.2021).

3.   Farhadi S., Asl H. D., Talebi Z. Morphological Awareness. Iran, 2015. Vol. 5, issue 4. 224 p. URL: https://mjltm.org/article-1-58-en.pdf (accessed: 17.05.2021).

4.   Gabbrielli M., Martini S. Programming Languages: Principles and Paradigms, Undergraduate Topics in Computer Science. 2010. URL: http://websrv. dthu.edu.vn/attachments/newsevents/content2415/Programming_Languages_-_Principles_and_Paradigms_thereds1106.pdf (accessed: 27.07.2021).

5.   Gee J. P. Unified discourse analysis: Language, reality, virtual worlds, and video games. New York: Routledge. 2014. P. 134–147.

6.   Herring S. Computer-mediated discourse analysis: an approach to researching online communities. Designing for Virtual Communities in the Service of Learning. 2004. P. 316–338.

7.   Jokuza E. Beyond dimensions: The man who married a hologram. CNN Digital Expansion. 2017. URL: https://edition.cnn.com/2018/12/28/health/rise-of-digisexuals-intl/index.html (accessed: 10.09.2019).

8.   Lazebna N. English Language as Mediator of Human-Machine Communication. Mysore, India : PhDians along with Ambishpere; Academic and Medical Publishers ; Royal Book Publishing, 2021. 571 p.

9.   Lazebna N., Prykhodko A. Digital discourse of English languageacquisition. Journal of Language and Linguistic Studies, 2021. URL: https://www.jlls.org/index.php/jlls/article/view/2533/852 (accessed: 25.01.2021).

10.   Makhachashvilli R., Semenist I., Zatsnyi Y., Klymenko O. Digital Interoperability of Foreign Languages Education. DHW 2021: Digital Humanities Workshop, 2021. P. 148–155.

11.   Petukhova V., Gropp M., Schmidt A. The DBOX Corpus Collection of Spoken Human-Human and Human-Machine Dialogues. 2014. URL: https://www.researchgate.net/profile/Volha_Petukhova/publication/304625403_The_DBOX_Corpus_Collection_of_Spoken_Human-Human_and_Human-Machine_Dialogues/links/577523ce08ae4645d60ba568/The-DBOX-Corpus-Collection-of-Spoken-Human-Human-and-Human-Machine-Dialogues.pdf   (accessed: 25.07.2020).

12.   RecipeQA. 2019. URL: https://hucvl.github.io/recipeqa/ (accessed: 18.12.2019).

13.   Scratch. URL: https://scratch.org/career-guides (accessed: 16.07.2021).

14.   Squeak by Example 5.3. Black A., Ducasse S., Nierstrasz O., Pollet D. USA. 2020. 320 p.

15.   Squeak. URL: https://squeak.com (accessed: 02.02.2021).

16.   VanLeeuwenT. Multimodality. In The Routledge handbook of applied linguistics. Routledge, 2011. P. 688–702.

17.   Yan Z., Duan N., Bao J.-W. DocChat: An information retrieval approach for chatbot engines using unstructured documents. 2016. URL: https://doi.org/10.3390/app10093335 (accessed: 26.07.2021).

Authors' bio

Nataliia Lazebna, Privat Dozent, Dr. Habil. TEFL Metodology Department, New Philological Institute, Julius-Maximilians University of Wuerzburg, Germany, was born on May 15, 1985 in Zaporizhzhia, Ukraine. She defended her Habilitation Project at Zaporizhzhia National University, Zaporizhzhia, Ukraine, in 2021, in the Field of Germanic languages. She obtained degree of Habilitated Doctor of Philological Sciences in the field of Germanic Languages (Confirmed by Zentralstelle für ausländisches Bildungswesen, Bonn, 21/04/2022, with the following Berufliche Anerkennung: habilitierter Wissenschaftler). She obtained her Ph.D. in Philological Sciences in the field of Germanic Languages in 2013 at Donetsk National University, Ukraine. Participated in AE E-Teacher Programs. TESOL Methodology certified (March, 2019). Alumni Cascade OPEN Courses with Colleagues (August - November, 2020). AE E-Teacher TESOL Methodology MOOC Facilitator at University of Maryland Baltimore County (February, 2020). For 15 years she has been working as Associate Professor at Department of Theory and Practice of Translation, Zaporizhzhia National Polytechnic University. The author of more than 50 academic journal articles, conference proceedings, and students' guidelines. The Official Opponent at 8 thesis defenses. The reviewer of vocabularies, monographs, and other printed scientific materials. An experienced translator and academic writer in the American and UK Companies.

Anatoliy Prykhodko, Doctor Habil. in Philology, Full Professor and Chair of Translation Theory and Practice Department at Zaporizhzhia National Polytechnic University, was born in Zaporizhzhia region, Ukraine, on 23 September, 1954. He graduated from Dniepropetrovsk State University in 1976. He got his Ph.D. degree in Philology, Germanic languages, in Kyiv State Pedagogical University of Foreign languages in 1982. He defended his postdoc thesis in 2002, Kyiv Shevchenko National University.
     Starting from 1976, he occupied various positions in Dniepropetrovsk, Kyiv and Zaporizhzhia. He has been delivering lectures in linguistics in Dniepropetrovsk, Kyiv, and Zaporizhzhia and passed his professional training at the universities of Germany (Essen, Han-

nover, Lepzig), including DAAD, Goethe Institute. He now is the Chair of Translation Theory and Practice Department at Zaporizhzhia National Polytechnic University.

Sphere of interest: syntax of Germanic languages, structural linguistic studies, cognitive linguistics, conceptology, linguosynergy, discoursive studies.

Dr. Habil. Prykhodko is a Member of All-Ukrainian Associations of Germanists.