

*Dovhan O. V.,  
PhD in Philology,  
Doctoral student at the Department of Slavic, Romance and Oriental Languages  
Drahomanov Ukrainian State University*

## LINGO-PRAGMATIC FEATURES OF NEURAL NETWORK MODELING OF LANGUAGE UNITS

**Summary.** The article highlights the place of neural network modeling of language units as a relevant and innovative tool for modern research (primarily, comparative linguistics, corpus, computer, etc. linguistics). Attention is focused on the productivity of the latter for use in the learning process: semantics (this refers to the ability of artificial neural networks to capture the context and peculiarities of transformational changes in the meaning of language corpora, which occurs due to properly structured training); lexicology (we are talking about modifications of the meaning of language units, the tracking of the genesis of which is possible due to the interaction of an artificial neural network with an array of text data (large corpora): thus, in the process of interaction, shades of meanings and changes in the actualization of certain lexical units (words) in different situations are taken into account); stylistics (this means improving the localization of complex sense constructions related to emotionality (irony, sarcasm, absurdity, etc.), which, in turn, will allow taking into account not only the surface meanings of words but also the specifics of their emotional differentiation, distinguishing, with the help of an artificial neural network, shades of semantics and their stylistic functionality); translation studies (the possibility of a structured representation of contextual data, which, in turn, produces a better distinction between shades of meaning and features of linguistic pragmatics related to the specifics of their actualization); integrated linguistic analysis (we are talking about cross-modal analysis, which studies intertextual, inter-auditory, and intermedial relations, which produces a deeper understanding of the peculiarities of the use of language units in the communication plane (ranked by several communicative acts and the specifics of their deployment and dynamics)), etc.

Thus, neural network modeling of language units actualizes innovative tools and produces new challenges and opportunities for industry research. We are talking about the prospects of improving neural network models on large data corpora (for example, lexicology, phraseology), building their training process (for example, semantics, morphology), and producing on its basis an understanding of complex language structures (for example, derivatology, syntax) and improving interaction in the context of language pragmatics (for example, semantics, stylistics). The above shows that neural network modeling of language units is a productive area of linguistic research. It is noteworthy that the impact of its existence in the context of technology and socio-cultural transformations is significant. First of all, it means that the study of linguistic and pragmatic aspects in this context will produce the creation of more intellectually intensive and adaptive systems at the intersection of the linguistic poly system and digitalization processes (for example, data science).

**Key words:** language units, language units analysis, text analysis, machine learning, artificial neural networks, neural network modeling.

**Statement of the problem in general terms and its connection with important scientific or practical tasks.** The process of neural network modeling, which is based on artificial neural networks that are similar in structure to the human brain, is a field that has great potential for modifying modern science. We are talking, first of all, about the humanities, since neural network modeling allows: a) processing of a large amount of data (for example, analysis of textual data), which, in turn, shows the effectiveness of their use for the study of discourse, discursive practices, etc; b) actualizing the context in the above-mentioned process: an artificial neural network can detect correlations in a large array of data (for example, large language models), representing them in all the complexity of mutual influence and building patterns, highlighting trends, etc; c) emphasizing the globalization component of the existence of modern science: the use of the above-mentioned technologies in the process of recognizing certain language structures, etc., and automated translation systems (for example, DeepL), etc, produces intensification of communication between representatives of different ethnic groups and exchange of experience between them, etc; d) identifying trends and patterns, to highlight possible integrated (intertextual, inter-auditory, intermedial) connections in a range of humanities studies, which produces the identification of the core themes of the latter and the formation of prospects for the development of the field, as well as the generation of ideas based on the collected data and within the framework of current concepts.

Thus, for the humanities, neural network modeling is an innovative tool with a fundamentally new research methodology, thanks to which the latter is being improved, complicated, and accelerated. First of all, this is due to the special, integrated nature of the artificial neural network construct, which underlies the above process. It is, in turn, a combination of computing power with creativity, knowledge intensity, and intellectuality that produces new approaches to the study and understanding of complex objects of analysis. Therefore, the humanities, by integrating neural network modeling into their paradigm, make a significant contribution to the development of understanding of the modern world in the context of the dynamics of socio-cultural, axio-gnosological, and other processes.

Therefore, the use of neural network modeling of language units is a promising area of research that is productive in the context of expanding the scope of application, taking into account linguistic differences in correlation with actualized cognitive processes, etc. In addition, the above-mentioned process is representative of the nature of the language poly system and the peculiarities of its variability (in particular, in the context of recognizing linguistic markers of the categories of sense and absurdity in political Internet discourse), which is the relevance of this research.

**Analysis of the latest research and publications on the topic, highlighting the previously unresolved parts of the general problem to which this article is devoted.** The problem of linguistic and pragmatic features of neural network modeling of language units has a multilayered nature, which, first of all, produces its actualization in several interdisciplinary and integrated works. Thus, in the study of A. Bonde et al. [1] highlight the linguistic and pragmatic features of actualizing natural language processing (hereinafter – NLP) to create query schemes for postoperative superficial infections. The authors emphasize that they have trained a universal language model of forward and backward reading on unlabeled postoperative cards. The scientists note that the above models were retrained on labeled data for SSSI classification, and the researchers propose the training implementation as an autonomous machine learning (SAM) and human-intervention (HIT) pipeline.

The experience of working with a multidimensional deep convolutional neural network with a long short-term memory with an attention mechanism (AM-MDC-LSTM) is presented in the study of H. Xu et al. [2]. In the analyzed study, the authors emphasize that updating the attention mechanism and using convolutional weights of different sizes significantly improves the performance of the neural network model. In particular, the researchers note that the above improves the work with individual exogenous variables and the extraction of local spikes and global periodic features with an obvious pattern. The authors consider it productive to combine these features with long short-term memory networks to extract temporal features reflected in certain groups of data.

A study of sentiment analysis as an important component of neural network modeling is presented in the study of H. Murfi et al. [3]. In this study, the authors present the specifics of sentiment analysis, a computer-based study of thoughts and emotions expressed in a text. The researchers note that in the process of sentiment analysis, the input text data is usually converted into a numerical representation. According to the researchers, the method of exact embedding is usually used for such a representation, which does not take into account the context of each word in the sentence. The authors see the solution to this shortcoming in updating the bidirectional encoding with transformation (BERT) model. The latter allows us to obtain textual representations based on the context and position of words in a sentence. In their work, the researchers extend the previous hybrid deep learning using the BERT representation to analyze Indonesian sentiment.

The use of transformational neural network models, such as GPT, which generate human-like speech and predict the human brain's reactions to language, is discussed in the study of G. Tuckute et al. [4]. In the analyzed study, the authors use brain responses to 1000 different sentences measured by functional MRI. By using a neural network model to identify new sentences, the researchers show the ability to predict the magnitude of brain responses associated with them. In addition, the authors use the neural network model to identify new sentences that are predicted to stimulate or inhibit responses in the human language network. The researchers emphasize that these new sentences, selected using the above model, do indeed strongly stimulate and inhibit the activity of human language areas in new people. The authors note that a systematic analysis of the sentences selected by the neural network model shows that the unexpectedness and quality of the formed construct on the linguistic input are the pivotal factors of the strength of the reaction in the language network. In turn, the results obtained by the research-

ers confirm the ability of neural network models not only to imitate human speech but also to non-invasively monitor neural activity in higher cortical areas, such as the language network.

The problem of the difficulties faced by neural network models in the process of executing and evaluating a small amount of training data is addressed in the study of Z. Wang et al. [5] In the analyzed study, the authors confirm the performance of neural network modeling of language units using multilayer neural networks. The scientists emphasize that the above networks are representative of performing similarity/difference search tasks with limited prior information. The researchers highlight the peculiarities of using multilayer neural networks (Siamese and triplet) for new machine learning systems. The authors present Siamese and triplet networks, focusing on their training and output processes. The researchers focus on hardware-efficient schemes for implementing the above networks. In addition, the researchers analyze and evaluate their effective training under different types of hardware errors, and present the latest noise-resistant methods that should be updated in working with them.

The correlation between the success of training a neural network model and the features of its architecture (in particular, the specifics of the latter's search) is presented in the study of R. Miikkulainen et al. [6]. In the analyzed study, the authors present the automated method CoDeepNEAT, designed to optimize deep learning architectures through evolution. The latter is achieved by extending existing neuroevolution methods to topology, components, and hyperparameters, which is productive for object recognition and language modeling. In addition, the aforementioned method supports the creation of a real-world application of automated image captioning on the journal's website. The researchers note that the expected growth in available computing power suggests that deep network evolution is a promising approach to building deep learning applications in the future.

The analysis of recurrent neural networks (hereinafter – RNNs), which are efficient for processing complex tasks (automatic translation, speech recognition, etc.), is presented in the study of D. Hopfe, K. Lee, C. Yu [7]. In this study, the authors examine various neural network models (RNNs, LSTM, GRU, deep LSTM, bidirectional LSTM, multidimensional RNNs, and multidimensional LSTM) in the context of efficiency and adaptability, comparing them with standard time series models (ARIMA, SARIMA, and SARIMAX) for short-term forecasting. The researchers argue that the inherent nonlinearity actualization of RNNs increases during unstable conditions, but it practically disappears during stable ones.

The specifics of the internal dynamics of an artificial neural network in the context of NLP (in particular, the peculiarities of understanding N400 sentences) are presented in the study of A. Lopopolo, M. Rabovsky [8]. In this study, the authors directly model the N400 amplitude obtained during the processing of naturalistic sentences, using as a predictor the update of the distributed representation of the sentence meaning generated by the sentence gestalt model trained on a large text corpus. The above, according to the scientists, makes it possible to quantitatively predict N400 amplitudes based on a cognitively motivated model, as well as to quantitatively compare this model with alternative N400 models. In particular, they compare the actualization measure from the SG model with the surprise measure estimated by a comparable language model trained on next-word prediction. The researchers argue that the results suggest that both sentence gestalt updating

and surprise predict aspects of N400 amplitude. Thus, they note that N400 amplitudes may reflect two distinct, but likely closely related, subprocesses that contribute to sentence processing.

Instead, the study of the correlation between deep learning algorithms for speech processing and the parameterization of pronunciation error detection and diagnosis is presented M. Lounis, B. Dendani, H. Bahi [9]. The analyzed study presents a thorough statistical analysis that the authors conducted by extracting specific data from 53 articles published in 2015–2023. The researchers note that their review shows that the diagnosis of pronunciation errors is a very active field of research. Many deep learning models and approaches have been proposed, but there are still some important open questions and limitations.

The specifics of linguistic data processing for Text-to-Speech (hereinafter – TTS) conversion are covered in the study R. Liu et al. [10], in which the authors note that with the development of deep learning, TTS models based on encoder-decoder demonstrate excellent performance. The researchers collected large unsupervised text data to pre-train a BERT-like language model and then applied the trained language model to extract deep linguistic information for the input text of the TTS model to improve the naturalness of the final synthesized speech. The researchers emphasize that to fully utilize the linguistic information associated with prosody in agglutinative languages, they incorporated morphological information into the training of the language model and built a morphology-based masking BERT model (MAM-BERT).

A theoretical and applied study of the linguistic means of expressing the category of temporality in modern political discourse is presented in the study D. Marieiev et al. [11]. In the analyzed study, the authors argue that the emergence of new challenges of today necessitates the modernization of approaches to understanding the category of time and its connection with the political life of the state. The scholars argue that the above results in the widespread use of linguistic means since the current trends of globalization require an adequate perception of the need for political change and its temporality. The researchers emphasize that the study of the phenomenon of temporality in political discourse as a two-pronged objective and subjective category of cognition of the world and political life in time, as well as the linguistic means of its expression, is becoming increasingly relevant. Thus, the study found that the category of temporality involves comprehension and interpretation of past political events, objective assessment and vision of the current state of functioning of the political sphere, and making the right decisions in the future. In addition, it has been found that the most common linguistic means of expressing temporality in modern political discourse are manipulations. The latter is an effective means of influencing the public consciousness and forming an opinion about the current state and trends in the development of political processes.

Thus, the analysis of the historiography on the problem of research on the lingo-pragmatic features of neural network modeling of language units has made it possible to highlight the prevailing trends in the modern scientific discourse and to highlight the gaps in it. Thus, despite the leading role of neural network modeling of language units as an innovative tool for humanities research (in particular, linguistic research), the analyzed historiography shows insufficient study of the actual linguis-

tic and pragmatic features of this process (for example, in the context of machine learning, etc.). Our research aims to fill these gaps and study the peculiarities of the above process in the context of working with textual data.

**Formation of the purpose of the article (statement of the task).** The purpose of the article is to consider the lingo-pragmatic features of neural network modeling of language units. The subject is the specifics of the above phenomenon in the context of machine learning and the work of artificial neural networks as an innovative tool of linguistic science.

**Presentation of the main research material with full justification of the scientific results obtained.** The above statement of the problem and the analysis of recent research and publications have shown the relevance of neural network modeling of language units in the context of the humanities (linguistics, philosophy of language, etc.) and natural sciences (physics, biology, etc.), as well as the sciences that are at their intersection (data science, statistics, etc.). First and foremost, the productivity of such modeling is related to the processes of language analysis and generation, which produce a more thorough understanding of its phenomenon, structure, parameterization, etc., as well as the genesis of language technologies.

However, for this study, it seems logical to start with the concept of “modeling”, since it is the core concept for understanding the specifics of its course and dynamics. For example, the Dictionary of the Ukrainian Language defines it through the action of “modeling”, representing it as “creating a model of someone or something” [12]), while the same source positions the concept of “model” as “a scheme of an object or phenomenon” [13]. Thus, modeling, in this interpretation, is positioned as the process of creating a model of someone or something that is inherently representative of the above.

It is noteworthy that the Cambridge Dictionary defines “modeling” as an activity “aimed at updating mathematical models (descriptions of a system or process) to predict or make certain calculations” [14]. The above concept must be close in meaning to “simulation”, which, in turn, also refers to the process of modeling. This is because it consists in creating “models of a set of problems or events that can be used to teach someone something”, or is, in fact, “the process of creating such a model” [15]. At the same time, under the concept of “model”, the source positions “something that can be copied because it is an extremely successful example of its type” [16]. Therefore, the process of modeling, based on the above, is an activity of updating models (representative examples of a certain type) for forecasting or making certain calculations.

Thus, modeling is an activity aimed at studying certain objects through the study of models representative of their functioning. The purpose of the modeling process is to comprehend the specifics of the latter's functioning and to predict their dynamics. It is noteworthy that the Cambridge Dictionary provides a more specific interpretation of the concept of the tasks of language modeling, which, in a narrow sense, is to predict the next word in a text based on the previous one. The result of such a process in the field of NLP is a specific practical application in the form of intelligent keyboards (for example, T9), suggestions of templates for possible responses to messages or e-mails (for example, several applications, applications, and browser extensions based on ChatGPT), automatic error correction (for example, Grammarly), etc.

The first language neural network model was built based on Yoshua Bengio's artificial neural network for direct propagation [17] in 2001. The input data for the aforementioned language neural network model were modified texts. We are talking about vectorization of the latter data – a process that consists of the numerical representation (by vectors of real numbers) of the input information so that the neural network model can process it. Today, the most popular variant of vector representations of words is word embeddings. The latter are compressed vectors that are transferred to a hidden layer, the output of which is then sent to the softmax layer, where the activation function is triggered. It is this function that is responsible for determining which of the signals will be further developed and which will not and is essentially responsible for the output of the entire model [18].

It is noteworthy that neural network modeling of language units is the basis for many further steps related to machine and deep learning. In particular, we are talking about the operation of the following neural network models: a) word2vec, which simplifies the process of language modeling; b) sequence-to-sequence (seq2seq), which generates the output sequence by predicting one word at a time: it is used for machine translation, data mining, etc.; c) pre-trained language models, which are used to represent language models for transfer learning, i.e., using the results of one artificial neural network to train another [18].

It is advisable to focus on the vector representation of words, which was further developed after the innovation of T. Mikolov, W.-t. Yih, G. Zweig [19], the essence of which is to remove the hidden layer and approximate the target. The aforementioned changes resulted in a more efficient implementation of the word2vec model, producing massive training of vector representations of words. It is noteworthy that the latter has two types that differ in their purpose: a) a “continuous bag of words” that predicts a central word based on the surrounding words, and b) a “skip-gram” that performs the opposite of the above type.

It should be noted that, in terms of their performance, the above vector representations do not differ from those acquired by an artificial neural network with a direct connection. However, training on a large corpus of texts allows them, unlike the above, to learn to approximate gender and species relations between words and their forms. It is noteworthy that the senses underlying such relations were, in fact, the primary interest of scientists (linguists, mathematicians, and others) in the vector representation of words. In particular, there are several studies devoted to the existence, dynamics, and transformational changes of the above-mentioned linear relations. However, subsequent works have shown the fluctuating nature of such relations: for example, the relations between words cannot always be determined objectively [18].

The aforementioned specificity has led to the development of vector representations of words, which has naturally become a core area of NLP. The latter is because updating pre-trained representations for the initialization process improves the performance of several tasks. For example, the relationship between words and their vector representations used to be largely unclear and almost mystical. However, the basis for the word2vec algorithm, which is based on determining such relationships, is matrix factorization methods.

In particular, we are talking about the use of classical approaches to matrix factorization: a) singular-value decomposition, which is a generalization of the eigenvalue decomposition of a matrix

of a nonnegatively definite normal matrix into an  $m \times n$  matrix as a generalization of the polar decomposition; b) latent semantic analysis, which is a method of processing information in natural language (in particular, distributional semantics) that highlights the correlation between a set of documents and the terms that are present in them by creating a set of concepts: the aforementioned method is based on the hypothesis that words with similar meanings will be present in similar texts. Thus, the use of the above approaches allows us to obtain the same results as with word2vec.

It is worth noting that the above algorithm is promising beyond the word level: for example, the skip-gram with a negative sampling model is productive. This is because such a model is convenient for training vector representations based on the local context, as well as for creating vector representations of sentences [19]. In addition, it is used outside the NLP domain: for example, in networks or biological sequences, etc. An interesting area is the projection of vector representations of words from different languages into a single space to provide interlanguage translation (from scratch). Teaching the latter in good quality exclusively without a teacher is becoming increasingly popular (at least for similar languages). The above-mentioned finds its application in translation from and to low-resource languages, as well as in machine translation without a teacher [18].

Over time, the aforementioned direct-coupled artificial neural networks have been replaced in language modeling by RNNs and long short-term memory artificial neural networks (LSTMs) [7]. It is noteworthy that the latter have undergone several transformational changes in recent years with the help of new language models that have significantly expanded their specific capabilities. It is worth noting that the above-mentioned RNNs are the most productive for language modeling since building a language model is a type of unsupervised learning. The latter allowed Yann LeCun [17] to call such learning predictive or predictive, which, in turn, led the researcher to position it as a core condition for the formation of strong artificial intelligence.

Subsequently, neural network models were widely expanded into NLP: it is noteworthy that three main types of artificial neural networks are most commonly used. We are talking about actualization:

1. RNNs are one of the most productive types due to their specific feature of having an internal memory. In addition, their significant advantage is the processing of huge amounts of data, which is optimal for linguistic research, as well as the emergence of the aforementioned long-term short-term memory (LSTM). In turn, the latter allows the user to update important details of the input data in the work of an artificial neural network, which would be lost in the process of their operation in other types of networks. This type of artificial neural network is evident when working with dynamic input sequences, which are ubiquitous in NLP. The aforementioned emergence of LSTM networks allowed us to eliminate the problem of the vanishing (or exploding) gradient.

2. Convolutional neural network (hereinafter – CNNs) – a powerful tool that updates a sequence of convolutional or merging layers to extract information from an image. This type of artificial neural network is productive for solving computer vision and image processing tasks. In the process of this neural network operation, convolutional layers are used to extract several elements (edges, textures, parts of objects). The above extraction process is performed by convolving the input data (image) using weights (a set of filters).

The latter is superimposed on the input data to create a map: this is done by calculating the dot product of the weights and the localized area of the image updated by the software in the process.

3. Recursive neural networks, the essence of which is the recursive application of the same set of weights to a structure. This type is used for structural or scalar prediction of input data (structures of variable size). The aforementioned process occurs by traversing a given structure in a certain sequence (topological): they are productive in learning sequences, which are representative of NLP based on the vector representation of words. In particular, they have proven to be an effective tool for learning distributed representations of structure – logic terms [20].

The real breakthrough in neural network modeling of language units was multitasking, a general approach in which neural network models are trained to perform various tasks on the same parameters. This is realized by linking their weights from different layers. In this way, the neural network model is trained to create a data representation that allows performing many tasks at once. It should be noted that sharing the same vector representations of words produces interaction and exchange of common low-level information – we are talking about “specific” representations of text elements.

The idea of multitasking learning was developed by R. Collobert, K. Kavukcuoglu, C. Farabet [21], whose work also put forward other ideas. In particular, the creation of pre-trained matrices of vector representations of words and the use of artificial CNNs to work with text data. The novelty of the latter was that this type of artificial neural network had not been actualized in NLP. In the context of the latter, this type of training is used in several cases: for example, a neural network model can be tasked with searching for mentions of inanimate objects in the analyzed data.

It is clear that developers do not need this information, but rather use it to clarify the main task: for example, it can be a prediction of the next word. First and foremost, updating the multitasking learning means that the main model will get better features at the input. In turn, evaluating the performance of neural network models based on the results of solving several tasks allows us to find out their generalization capabilities.

It is noteworthy that the vast majority of achievements in the field of NLP can be reduced to one type of neural network modeling of language units. However, to achieve a real breakthrough and the cornerstone of integrated artificial intelligence research – natural language understanding – simple training on raw, unprocessed texts is not enough. This is because it is necessary to search for new methods and create neural network models with different functionality. However, the approach to language modeling based on the above-mentioned RNNs is currently dominant, due to the ability of this type of artificial neural network to process unlimited context [22].

**Conclusions from the study and prospects for further research in this area.** In the context of linguistic science, neural network modeling is a relevant and innovative tool for conducting modern research (first of all, we are talking about comparative linguistics, corpus, computer, etc. linguistics). In particular, neural network modeling is productive for use in the learning process:

– *semantics* (this refers to the ability of artificial neural networks to capture the context and peculiarities of transformational changes in the meaning of language corpora, which occurs due to properly structured training);

– *lexicology* (we are talking about modifications of the meaning of language units, the tracking of the genesis of which is possible due to the interaction of an artificial neural network with an array of text data (large corpora): thus, in the process of interaction, shades of meanings and changes in the actualization of certain lexical units (words) in different situations are taken into account);

– *stylistics* (this means improving the localization of complex sense constructions related to emotionality (irony, sarcasm, absurdity, etc.), which, in turn, will allow taking into account not only the surface meanings of words but also the specifics of their emotional differentiation, distinguishing, with the help of an artificial neural network, shades of semantics and their stylistic functionality);

– *translation studies* (the possibility of a structured representation of contextual data, which, in turn, produces a better distinction between shades of meaning and features of linguistic pragmatics related to the specifics of their actualization);

– *integrated linguistic analysis* (we are talking about cross-modal analysis, which studies intertextual, inter-auditory, and intermedial relations, which produces a deeper understanding of the peculiarities of the use of language units in the communication plane (ranked by several communicative acts and the specifics of their deployment and dynamics)), etc.

Thus, neural network modeling of language units actualizes innovative tools and produces new challenges and opportunities for industry research. We are talking about the prospects of improving neural network models on large data corpora (for example, lexicology, phraseology), building their training process (for example, semantics, morphology), and producing on its basis an understanding of complex language structures (for example, derivatology, syntax) and improving interaction in the context of language pragmatics (for example, semantics, stylistics).

The above shows that neural network modeling of language units is a productive area of linguistic research. It is noteworthy that the impact of its existence in the context of technology and socio-cultural transformations is significant. First of all, it means that the study of linguistic and pragmatic aspects in this context will produce the creation of more intellectually intensive and adaptive systems at the intersection of the linguistic poly system and digitalization processes (for example, data science).

#### **Bibliography:**

1. Assessing the utility of deep neural networks in detecting superficial surgical site infections from free text electronic health record data / A. Bonde et al. *Frontiers in Digital Health*. 2023. Volume 5. <https://doi.org/10.3389/fdgth.2023.1249835> *Frontiers* : website. URL: <https://goo.su/Xza8a> (date of application: 26.01.24).
2. Attention Mechanism Multi-size Depthwise Convolutional Long Short-term Memory Neural Network for Forecasting Real-time Electricity Prices / H. Xu et al. *IEEE Transactions on Power Systems*. 2024. <https://doi.org/10.1109/TPWRS.2024.3353759> *IEEE Xplore* : website. URL: <https://goo.su/UafzF3p> (date of application: 26.01.24).
3. BERT-based combination of convolutional and recurrent neural network for Indonesian sentiment analysis / H. Murfi et al. *Applied Soft Computing*. 2024. Volume 151. <https://doi.org/10.1016/j.asoc.2023.111112> *ScienceDirect* : website. URL: <https://goo.su/AJbo5D0> (date of application: 26.01.24).
4. Driving and suppressing the human language network using large language models / G. Tuckute et al. *Nature Human Behaviour*. 2024. <https://doi.org/10.1038/s41562-023-01783-7> *Nature Human Behaviour* : website. URL: <https://goo.su/OQGS3> (date of application: 26.01.24).

5. Emerging Machine Learning Using Siamese and Triplet Neural Networks / Z. Wang et al. Design and Applications of Emerging Computer Systems. Cham: Springer Nature Switzerland, 2024. P. 115–141. [https://doi.org/10.1007/978-3-031-42478-6\\_5](https://doi.org/10.1007/978-3-031-42478-6_5) Springer Link : website. URL: <https://goo.su/nZaVW> (date of application: 26.01.24).
6. Evolving deep neural networks / R. Miikkulainen et al. Artificial intelligence in the age of neural networks and brain computing. 2024. P. 269–287. <https://doi.org/10.1016/B978-0-323-96104-2.00002-6> ScienceDirect : website. URL: <https://goo.su/zjGzv0e> (date of application: 26.01.24).
7. Hopfe D. H., Lee K., Yu C. Short-term forecasting airport passenger flow during periods of volatility : Comparative investigation of time series vs. neural network models. Journal of Air Transport Management. 2024. Volume 115. <https://doi.org/10.1016/j.jairtraman.2023.102525> ScienceDirect : website. URL: <https://goo.su/YpALKh> (date of application: 26.01.24).
8. Lopopolo A., Rabovsky M. Tracking lexical and semantic prediction error underlying the N400 using artificial neural network models of sentence processing. Neurobiology of Language. 2024. P. 1–69. [https://doi.org/10.1162/nol\\_a\\_00134](https://doi.org/10.1162/nol_a_00134) MIT Press Direct : website. URL: <https://goo.su/n51tTJ> (date of application: 26.01.24).
9. Lounis M., Dendani B., Bahi H. Mispronunciation detection and diagnosis using deep neural networks: a systematic review. Multimedia Tools and Applications. 2024. P. 1–35. <https://doi.org/10.1007/s11042-023-17899-x> Springer Link : website. URL: <https://goo.su/SIbX6> (date of application: 26.01.24).
10. Text-to-Speech for Low-Resource Agglutinative Language With Morphology-Aware Language Model Pre-Training / R. Liu et al. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2024. Volume 32. P. 1075–1087. <https://doi.org/10.1109/TASLP.2023.3348762> IEEE Xplore : website. URL: <https://goo.su/s12H> (date of application: 26.01.24).
11. Linguistic Means of Expressing the Category of Temporality in Modern Political Discourse / D. Mariciev et al. World Journal of English Language. 2023. Volume 13, Issue 4. P. 1–23. ResearchGate : website. URL: <https://goo.su/qo9ZDPn> (date of application: 26.01.24).
12. Моделювати. Горох : вебсайт. URL: <https://goo.su/NcZfdPU> (дата звернення: 26.01.24).
13. Модель. Горох : вебсайт. URL: <https://goo.su/FHWY75F> (дата звернення: 26.01.24).
14. Modeling. Cambridge Dictionary : website. URL: <https://goo.su/mnYaJoW> (date of application: 26.01.24).
15. Simulation. Cambridge Dictionary : website. URL: <https://goo.su/0KFxF0a> (date of application: 26.01.24).
16. Model. Cambridge Dictionary : website. URL: <https://goo.su/2uAszx> (date of application: 26.01.24).
17. LeCun Y., Bengio Y., Hinton G. Deep learning. Nature. 2015. Volume 521. P. 436–445. <https://doi.org/10.1038/nature14539> Creative Coding : website. URL: <https://goo.su/pbyX> (date of application: 26.01.24).
18. Ruder S. A Review of the Neural History of Natural Language Processing. QUANTEXA : website. 2018. URL: <https://goo.su/n8QhbAq> (date of application: 26.01.24).
19. Mikolov T., Yih W.-t., Zweig G. Linguistic Regularities in Continuous Space Word Representations. Proceedings of NAACL-HLT 2013. Atlanta : Georgia, 2013. P. 746–751. ACL Anthology : website. URL: <https://goo.su/nKCe3> (date of application: 26.01.24).
20. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank / R. Socher et al. The Stanford Nature Language Processing Group : website. 2013. URL: <https://goo.su/MiFaZz> (date of application: 26.01.24).
21. Collobert R., Kavukcuoglu K., Farabet C. Torch7 : A Matlab-like Environment for Machine Learning. EPFL : website. 2011. URL: <https://goo.su/DE2PI> (date of application: 26.01.24).
22. Language Modeling with Gated Convolutional Networks / Y. N. Dauphin et al. ARXIV : website. 2017. URL: <https://goo.su/wbaxzs4> (date of application: 26.01.24).

### Довгань О. Лінгвопрагматичні особливості нейромережевого моделювання мовних одиниць

**Анотація.** Стаття висвітлює місце нейромережевого моделювання мовних одиниць як актуального та інноваційного інструменту для проведення сучасних досліджень (першою чергою, мовиться про зіставне мовознавство, корпусну, комп'ютерну тощо лінгвістику). Акцентовано увагу на продуктивності останнього для використання в процесі вивчення: семантики (йдеться про здатність штучних нейронних мереж вловлювати контекст та особливості трансформаційних змін значення мовних корпусів, що відбувається за рахунок правильно вибудованого навчання); лексикології (мовиться про видозміни значення мовних одиниць, відстеження генези якого можливе завдяки взаємодії штучної нейронної мережі з масивом текстових даних (великих корпусів): так, в процесі взаємодії відбувається врахування відтінків значень та змін у актуалізації тих чи тих лексичних одиниць (слів) у різних ситуаціях); стилістики (мається на увазі вдосконалення локалізації складних смислових конструкцій, пов'язаних з емотивністю (іронія, сарказм, абсурдизація тощо), що своєю чергою дозволить врахувати не лише поверхневі значення слів, а й специфіку їх емотивної диференціації, розрізняючи, за допомогою штучної нейронної мережі, відтінки семантики та їх стильову функційність); перекладознавства (можливість структурованої репрезентації контекстуальних даних, що, своєю чергою, продукує краще розрізнення відтінків значень та особливостей мовної прагматики, яка пов'язана з специфікою їх актуалізації); інтегрованого лінгвістичного аналізу (мовиться про кросмодальний аналіз, у межах якого вивчають інтертекстуальні, інтраудіальні та інтермедіальні зв'язки, що продукує глибше розуміння особливостей побутування мовних одиниць в комунікаційній площині (ранжовані низкою комунікативних актів та специфікою їх розгортання й динаміки)) тощо.

Отже, нейромережеве моделювання мовних одиниць актуалізує інноваційний інструментарій та продукує нові виклики й можливості для галузевих досліджень. Мовиться про перспективи вдосконалення нейромережевих моделей на великих корпусах даних (до прикладу, лексикологія, фразеологія), вибудовування процесу їх навчання (до прикладу, семантика, морфологія) та продукування на його підґрунті розуміння ними складних мовних конструкцій (до прикладу, дериватологія, синтаксис) й вдосконалення взаємодії в контексті мовної прагматики (до прикладу, семантика, стилістика). Вищезазначене засвідчує, що нейромережеве моделювання мовних одиниць є продуктивним напрямом лінгвістичного дослідження. Прикметно, що вплив його побутування в контексті функціонування технологій та соціокультурних трансформацій є вагомим. Мовиться, першою чергою, про те, що дослідження лінгвопрагматичних аспектів у означеному контексті продукуватиме створення інтелектуальноємніших та адаптивніших систем на перетині побутування мовної полісистеми й диджиталізаційних процесів (до прикладу, наука про дані).

**Ключові слова:** мовні одиниці, аналіз мовних одиниць, текстовий аналіз, машинне навчання, штучні нейронні мережі, нейромережеве моделювання.