Revenko Ye. S.
Kherson State University

FAKE NEWS ANALYSIS: LEXICAL AND SEMANTIC DECEPTION CUES

This article is devoted to the study of the lexical and semantic features of deception in modern media space. Language is a flexible and a most important tool for communication. It provides us with the opportunity to discover the world and share beliefs and opinions. Nowadays in the era of post-truth and information war the language, being misused, can become a dangerous weapon. Thus, in recent years many techniques and tools have been created to reduce the spread of disinformation. Fake news detection becomes one of the key tasks not only for professional journalists and researches but also for tech giants like Google and Facebook. This research concentrates on the linguistic analysis of fake news texts as one of the most reliable methods of disinformation detection in the media space. The review of scientific sources on the research topic is presented. The concepts "post-truth" and the categories of fake news are also considered. The categories of the fake news include serious fabrications, large-scale hoaxes and humorous fakes. Using the content analysis of fake news corpus, the specific thematic focus that is characteristic of fake news is distinguished in the article. The article also includes the characteristic of fake news headlines. The predictive function of the headlines, which points at the content of the text and a phatic function, which is to attract the attention of, and establishing the first contact with any prospective reader are outlined. The studied corpus of fake news makes it possible to identify such lexical and stylistic features of fake news as the informal non-objective narrative character; the use of metaphors and stylistically marked vocabulary; an abnormality at all levels of content (thematic, informative, lexical, syntactic); the use of trigger words and hedging words; the frequent use of adjectives; an expressly ironic or sarcastic tone. It is concluded that fake news is characterized by a wider emotive space than denotative one.
Key words: *post-truth, media discourse, disinformation, information war, clickbait.*

Formulation of the problem. The modern global communicative culture exists in the "post-truth" or "post-communication" era. The Oxford Dictionary defines "post-truth" as "circumstances in which people respond more to feelings and beliefs than to facts" [12]. Under such circumstances, the deception can spread through media space much faster than the truth since the deception appeals to the lesser, baser instincts of the recipients.

Words in the media have significant power in shaping beliefs and opinions. Due to the rapid development of technology and increasing opportunities for any user to share their opinions with the public, the truthfulness of the words is endangered. Thousands of fake news are daily disseminated in the media space to maximize manipulative or propaganda impact. That is why it is so important to detect objective linguistic cues that indicate that we are dealing with misinformation in order to improve existing fact-checking resources.

Thus, the analysis of the semantic and linguistic features of fake news is relevant for understanding the role of fake news in creating the global infor-

mation space, and therefore the language space as a whole, which according to the cognitive science is the most important category in the perception of the world by a person.

Each text of fake news has a subjective frame of intentional orientation of the author and is designed to cause a corresponding reaction by the recipient of the message. Linguistic techniques are functional tools for realizing the interests of the author by influencing the processes of cognitive reflection of the recipient.

The objective of the article is to analyze the semantic and linguistic characteristics of fake news and to develop a system of aspects which can help identify the deception.

1. Literature review

Fake news with their direct impact on the information environment is the sphere of scientific interests of both domestic and foreign researchers. The issue of fake news was investigated mainly in two directions: (1) distribution channels of fake news and (2) analysis of the fake news content.

The issue of the distribution of fake information in Ukrainian social media is outlined in the scientific

works of M. Kitsa [6] and I. Mudra [9]. The fake news in the linguistic-cognitive and synergetic aspects were studied by Yu. Hlavatska [5] and Yu. Omelchuk [11].

Linguistic aspects of deception detection were studied in a variety of works. Rubin, V., Chen, Y., & Conroy, N. in the work '*Deception detection for news: Three types of fakes*' [14] separated the task of fake news detection by type of fake: serious fabrications, large-scale hoaxes and humorous fakes. Niall J Conroy et al. in the research work '*Automatic deception detection: Methods for finding fake news*' [1] provided a typology of veracity assessment methods emerging from two major categories – linguistic cue approaches and network analysis ones.

The most relevant work to our research is '*Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking*' written by Hannah Rashkin et al [13]. This scientific work presents an analytic study on the language of news media in the context of political fact-checking and fake news detection. The authors compare the language of real news with that of satire, hoaxes, and propaganda to find linguistic characteristics of the untrustworthy text.

2. Categories and thematic focus of fake news

In order to address the fake news problem, it is necessary to begin with distinguishing of fake news types. At the present time, there is no universally agreed definition and classification of fake news in the scientific community. Fake news becomes an object of studies which are based on various factors such as news distribution channels, its goals, countermeasures and so on. The researchers distinguished two types of fake news according to genre characteristics: 1) fake news which have all the genre signs of real news and whose purpose is to deceive recipients; 2) fake news which is a parody of real news and aims at entertaining recipients rather than reporting information [2].

Victoria L. Rubin et el. define three categories of fake news with their key differences depending on the author's intention [14]. They are the following:

− *Serious fabrications* are examples of exposed fraudulent journalistic writing.

Yellow journalism, political propaganda, click-baits, conspiracy theories belong to this category. Serious fabrications deliberately highlight topics such as sensational crime stories, astrology, gossip columns about celebrities, and junk food news.

− *Large-Scale Hoaxes* is another type of fabrications which look like real news

but contain false information. Hoaxes attempt to deceive recipients pretending to be real news and often mistakenly reported by respected news agencies.

− *Humorous Fakes* (news satire, news parody) – the main communicative purpose is directed at the entertainment of the reader but not at the message of information. The key difference between humorous fake news and other types is that the authors use the means and techniques of real news text and at the same time signal to the reader that it is a parody rather than a statement of facts. In order for authors' intentions to be clear to readers, the text is saturated with various obvious markers of comical and events are described in absurd manner.

The specific thematic focus can be cues of deception. It is usually about controversial topics and tends to cause strong emotional reactions. Using contextual analysis of fake news corpus, the following topics typical of fake news can be distinguished [3]:

− hyperbolic statements against one person or group;

E.g. *Putin has stated: Russian citizens (have) 2 options: Stay home for 15 days or in jail for 5 years* (Source: Instagram post).

− hyperbolic statements in favor of one person or group;

E.g. *Nate McMurray wants to reward illegals with amnesty* (Source: YouTube, "Across," Jacobs for Congress).

− sensational crimes and violence;

E.g. *Texas man admits kidnapping 79 people* (Source: Facebook post).

− racist messaging;

Donald Trump said, "Immigrants aren't people, they're animals." (Source: Facebook post).

− paranormal theories;

E.g. *Alien discovery: Two 'bases' spotted on Mars which add weight to life on Mars theory* (Source: the online edition "Express").

− conspiracy theories.

E.g. *Why Did Beyonce Name Blue Ivy "Blue Ivy"? Ask The Illuminati* (Source: the online edition "Romper").

One of the most common themes of fake news is the hyperbolic criticism of politicians. Fake news raises awareness by referring to well-known public figures, usually Donald Trump, Barack Obama, Hillary Clinton and Vladimir Putin. This can be explained by the fact that fake news is directly related to the information warfare, the purpose of which is the distortion of the facts for political propaganda and manipulation of the mass consciousness as well as creating the image of the enemy. The second most common thematic focus of fake news is conspiracy theories. Fake news with a focus on conspiracy theories are characterized by specific vocabulary, for

example, the use of the phrases *"a secret government"*, *"transnational corporations"*, *"a world government"* and the verbs *"to deceive"*, *"to hide"*, *"to conceal"*. Representations of sensationalist crimes are also common for fake news. Paranormal theories are more apparent in satirical fake news.

3. Linguistic cues of deception

Each segment of linguistic impact on public consciousness has its own set of manipulative tricks and techniques. Researchers believe that fake and real news has different ways of presenting the information.

3.1. Characteristic of fake news headlines

The headline can be considered as an independent communication unit that is also a part of the following text. All headlines are meta texts identifying their co-text and establishing the first contact with an audience. Researchers distinguish the predictive function of the headlines (it points at the content of the text) and a phatic one (it attracts the attention and establishes the first contact with any prospective reader) [10].

Namely, the headlines should attract readers' attention instantly since on average there are more people reading only headlines than reading the whole article. Most of the headlines appeal to the emotions of the reader, intend to shock, scare or terrify them, so an impressed person is no longer able to evaluate information critically. The mechanism of action behind crime articles, pseudoscientific, political and other types of fake news is often based on curiosity, anxiety or sexual undertones.

The following headline illustrates the foregoing: *L.A.'s Elite Cannibal Restaurant Boasts Katy Perry, Meryl Streep, Chelsea Clinton As Members.* The headline intents to shock and horrify the readers by mentioning the cultural prohibition – cannibalism. Such a topic causes readers a feeling of disgust and at the same time curiosity. The mention of famous personalities encourages readers to click and share the fake news.

The headlines of fake articles are often hyperbolical, sensational, and striking, or rather "clickbating". The online dictionary by Merriam-Webster defines "a clickbait" as something (such as a headline) designed to make readers want to click on a hyperlink especially when the link leads to content of dubious value or interest [8].

In the following example, there is no evaluative vocabulary, but there is a trigger (the words that provoke a person to certain emotions or actions) word *'BREAKING'* that will force the reader to follow the link and read the material: *BREAKING: Illegal Muslim From Iran Arrested For Starting California Wildfire.* Besides, the author of this fake news appeals to religious conflicts using the expression "*Illegal Muslim*" and to tense relations between the United States and Iran, indicating that a person is from Iran.

Similar trigger words which can be found in headlines are *"warning"*, *"attention"*, *"you won`t believe"*, *"you will be shocked"*; they are often written in capital letters. For example, *Man Tries To Hug a Wild Lion, You Won`t Believe What Happens Next!; You'll Be SHOCKED By How Shredded His Abs Are After Eating This Superfood.*

Headlines formed in accordance with such a lexical and semantic construction are capable of producing expected effect even in the absence of other manipulative techniques and seemingly for the manipulator negative thematic content of the main part of the text. In the following example of the fake news the trigger words are not used but the semantic construction of the headline addresses reader emotions and arouse curiosity, implicitly pointing to the topic of pedophilia and death: *The bones of CHILDREN have been found on Epstein's island.*

The headline usually aims to provoke readers` emotions, thereby increasing the probability that the reader will share information through social media. Regarding this issue, Brooke Binkowski, an award-winning journalist and researcher, emphasizes that if after reading the headline, the reader has a feeling of fear, anger, or disappointment, then the news and source should be checked for deception.

3.2. Lexical and stylistic features of fake news

The emotive space of the fake news is noticeably wider than the denotative one. Fake news is distinguished by stylistically marked vocabulary and metaphorization, so the expression focus plays a much more important role than the content. While real news is created in order to convey information as objectively as possible, allowing the reader to evaluate events by themselves, the fake news is characterized by the informal non-objective narrative character and often giving a personal assessment of events by the author.

An example of the above statements is the fake news that was circulated in March 2020 throughout the media space: *This picture is from World War II, a soldier carrying a donkey. It is not that the soldier loves donkeys or has some sort of perversion. What's happening is that the field is mined and that if the donkey was free to wander as it pleased, it would likely detonate a charge and kill everyone. The moral of the story is that during difficult times the first ones you have to keep under control are the jackasses who don't understand the danger and do as they please.* This news was accompanied by a photograph of a soldier transporting a donkey on his back.

Fig. 1. Illustration for the fake news article about COVID-19

The author of the fake news gives a personal assessment of events metaphorically comparing World War II with the COVID-19 coronavirus pandemic and associating donkeys with people who can cause harm to others through their ignorance. The use of a non-literary word "*the jackasses*" and appeal to taboo topics "*It is not that the soldier loves donkeys or has some sort of perversion*" arouse interest among readers and indicate deception.

The correlation between norm and abnormality at all levels of content (thematic, informative, lexical and syntactic) may indicate falsity of news. Those facts or events that go beyond the norm rather attract the attention of the reader than ordinary things.

For example, the fake news which tells about the Russian Sleep Experiment has been distributed on many sites and social networks by thousands of users: *Russian researchers in the late 1940's kept five people awake for fifteen days using an experimental gas based stimulant. They were kept in a sealed environment to carefully monitor their oxygen intake so the gas didn't kill them, since it was toxic in high concentrations. This was before closed circuit cameras so they had only microphones and 5 inch thick glass porthole sized windows into the chamber to monitor them.* At first glance it seems that this is real news, since the language of this text is news-like, the author imitates the real press and the material is written in accordance with the canons of journalism. The deception cues, in this case, are the lack of a competent character system (the author does not indicate the persons` names who take part in the experiment), the uncertainty of the chronotope (the lack of exact dates) and abnormality of the described events. A hyperbolic description and a tense emotional background signalize that this is fake news.

The frequent use of adjectives in the fake news contributes to a clear visualization of the presented information and can affect the emotional perception of the news.

For example, the fake news article with the headline "*On This Day In 2014, Obama Tried To Change The Presidential Seal*" tells about the intentions of Barack Obama to change the design of the Presidential Seal: *On this date, 5 short years ago, President Obama tried to make a radical change to the look of the office he held. He attempted to change the Presidential Seal with a new one that reflected his massive ego and tiny brain...The only other addition in imagery was a "$" within the center circle, presumably symbolic of all the riches he intended to steal during his evil reign.* Even in such a short quotation, the expressive vocabulary which creates a dismissive connotation of the text can be distinguished: "*massive ego*", "*tiny brain*", "*evil reign*". The excessive use of adjectives is an explicit cue of deception.

The authors of real news use assertive words as well as concrete figures and comparatives which indicate that they are more accurate about describing events. Fake news authors often use hedging or also known as vague language. Vague language is not precise and therefore allows speakers not to commit themselves or not to sound too definite [7].They state a general idea but the precise meaning creates a recipient of information. The authors of fake news use hedging words such as "*assume*", "*believe*", "*suggest*", "*allegedly*" in order to make news articles less direct, and limit or qualify claims and statements. The statements they make can`t be proven, thus authors of fake news can protect themselves from criticism and potential legal action in the process. For example: *As Epstein Died, Guards <u>Allegedly </u>Shopped Online and Slept; Iran Says It Arrested 17 Iranians <u>Allegedly </u>Recruited by CIA; Marcus Bachmann Refused Service in Indiana, Store Owner <u>Assumed</u> He Was Gay.*

A reliable marker of deception is the lack of reference to the real source of information. Such expressions as "*an anonymous source*", "*reportedly*", "*witnesses say*", "*according to witnesses*" are often used in fake news. For example, <u>*Witnesses Say*</u> *Greg Gianforte Grabbed Guardian Reporter and Slammed Him to the Ground; Trump Viewed Ukraine as Adversary, Not Ally, <u>Witnesses Say</u>; Navy Chief <u>Reportedly </u>Beat Up Kid at Youth Soccer Game; iPhone FaceTime Bug <u>Reportedly </u>Allows Eavesdropping.*

The call to share should cause readers to doubt the veracity of the news. Reliable news sites are often well-known, so they do not need advertising and do not encourage readers to spread the news on the Internet. Most fake news urge users to share the articles implicitly or explicitly: *Please, share this to reach*

as many people as possible!; PLEASE, SHARE THIS WITH OTHERS!.

An expressly ironic or sarcastic tone of the news article indicates fake news with high probability. The essence of sarcasm realization is manifested in the grotesque, parody and absurdity of the situation, described the unfolding of the conflict: the "vivid" description of events in the text of the "fake" news is incompatible with real world events [4, p. 24]. Although satirical news sites such as *The Onion* position themselves as an entertaining source and do not correlate themselves with fake news, in some cases their news articles have been mistaken for actual reporting. For example, a lot of news editions spread a report by The Onion naming North Korean dictator Kim Jong-un as the sexiest man alive, not realizing it is a satire. The news editions cited the Onion: *"Blessed with an air of power that masks an unmistakable cute, cuddly side, Kim made this newspaper's editorial board swoon with his impeccable fashion sense, chic short hairstyle, and, of course, that famous smile. Photos with the story include Kim on horseback squinting into the light, Kim waving towards a military parade, Kim wearing sunglasses and smiling, and touring a facility with his wife."* In addition to obvious irony, the passage contains excessive use of the adjective, which also indicates fake news.

The unlimited functional potential of the language and the accessibility of the manipulative framework of communicative interaction determine the linguistic aspect as a key tool for latent influence on public consciousness, an effective mechanism for programming the archetypes of behavior and thinking of social masses.

Conclusions. We conclude that fake news is a global issue that affects not only the political sphere and the information war as a phenomenon of modern media discourse but also intervenes in other informational aspects such as climate change, legislation, entertainment and others.

Fake news is different from high-quality journalism that conforms to the professional standards and ethics but at the same time. Emotions in modern media space become more important than objectivity and reliability. Thus, the fake news is often more attractive to the audience than reliable information. In the context of globalization and technological development everyone can publish and share beliefs and opinions that are in the public domain. A huge number of alternative sources make it difficult to verify the information for reliability. Therefore, not only professional journalists, researchers and developers of social networks but even ordinary citizens set themselves the task of detecting fake news, creating new tools for analyzing and evaluating information content.

In our article we try to collect and classify mechanisms of fake news detection. We present the categories and thematic focus of fake news and identify linguistic cues of deception that will help readers to indicate fake news and increase their media literacy.

We consider the analysis of linguistic and cognitive specificity and communicative and pragmatic peculiarities of fake news to be our research perspective.

References:
1. Conroy N., Rubin V., Chen Y. Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology. 2015. 52(1):1–4. DOI: 10.1002/pra2.2015.145052010082.
2. Edson C., Zheng W. Defining "Fake News": A typology of scholarly definitions. *Digital Journalism.* 2017. DOI:10.1080/21670811.2017.1360143.
3. Golbeck J., Mauriello M., Auxier B., Bhanushali K., Bonk C., Bouzaghrane M., et al. 2018. Fake news vs satire: A dataset and analysis. In Proceedings of the 10th ACM Conference on Web Science, P. 17–21. DOI: 10.1145/3201064.3201100.
4. Hlavatska Yu. L. "Fake" news via the comic: distortion of information. Modern development trends in philology and social communication : collective monograph. 2019. P. 18–33.
5. Hlavatska Yu. L. Klasyfikatsiia feikovykh novyn u suchasnomu mediaprostori: synerhetychnyi aspekt [Classification of "fake" news in modern media space: synergrtic aspect]. *Germanic studies and intercultural communication*, 2019. № 1, P. 275–280.
6. Kitsa M. O. Feikovainformatsiia v ukrainskykh sotsialnykh media: poniattia, vydy, vplyv na audytoriiu [Fake information in Ukrainian social media" concepts, types, impact on audience]. *Scientific Notes. Social communications.* 2016. №. 1 (52), P. 281–286.
7. Macmillan online dictionary. 2020. URL: https://www.macmillandictionary.com/dictionary/british/vague-language (accessed: 09.04.2020).
8. Merriam-Webster online dictionary. 2020. URL: https://www.merriam-webster.com/dictionary/clickbait (accessed: 09.04.2020).

9. Mudra I. M. Poniattia "feik" ta yoho vydy u ZMI. *Tele- ta radiozhurnalistyka.* 2016. №. 15. P. 184–188.

10. Nord C. Text-Functions in Translation: Titles and Headings as a Case in Point. Target. *International Journal of Translation Studies.* 1995. Vol. 7, № 2, P. 261–284. DOI: https://doi.org/10.1075/target.7.2.05nor.

11. Omelchuk Yu. O. Pseudo News as a Genre of Modern English Media Discourse: Linguistic, Cognitive, Communicative and Pragmatic Parameters: Dissertation for PhD: 10.02.04 / Zaporizhzhia National University. Zaporizhzhia, 2018. 230 p.

12. Oxford Dictionary. Word of the year of 2016. URL: https://languages.oup.com/word-of-the-year/2016/ (accessed: 09.04.2020).

13. Rashkin H., Choi E., Jang J., Volkova S. Truth of varying shades: Analyzing language in fake news and political fact-checking. In Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics. 2017. P. 2931–2937. DOI: 10.18653/v1/D17-1317.

14. Rubin V., Chen Y., Conroy N. 2015. Deception detection for news: Three types of fakes. Proceedings of the Association for Information Science and Technology. 52(1):1–4. DOI: 10.1002/pra2.2015.145052010083.

Ревенко Є. С. АНАЛІЗ ФЕЙКОВИХ НОВИН: ЛЕКСИКО-СТИЛІСТИЧНІ МАРКЕРИ НЕПРАВДИ

Стаття присвячена дослідженню лексико-семантичних маркерів неправди в сучасному медіапросторі. Мова – це гнучкий і найважливіший інструмент спілкування. Вона дає нам можливість відкривати світ та ділитися переконаннями та думками. Сьогодні, в епоху постправди та інформаційної війни, мова може стати небезпечною зброєю у разі її неправильного використання. Враховуючи це, останніми роками було створено багато методик та інструментів для зменшення поширення дезінформації. Виявлення фейкових новин стає одним із ключових завдань не тільки для професійних журналістів та дослідників, але й для таких технічних гігантів, як Google ma Facebook. Наше дослідження зосереджене на лінгвістичному аналізі текстів фейкових новин як одному з найбільш надійних методів виявлення дезінформації в медіапросторі. У статті представлено огляд наявної літератури на тему дослідження. Також розглянуто поняття «пост правда» та окреслено категорії фейкових новин. До категорій фейкових новин належать серйозні фабрикації, масштабні містифікації та жартівливі фейки. За допомогою аналізу контенту корпусу фейкових новин встановлено специфічну тематичну спрямованість, характерну для фейкових новин. У статті наведено характеристику заголовків фейкових новин. Визначено прогностичну функцію заголовків, яка вказує на зміст тексту та фактичну функцію, що полягає у привабленні уваги та налагодженні першого контакту з будь-яким потенційним читачем. Вивчений корпус фейкових новин дає змогу виявити такі лексичні та стилістичні особливості фейкових новин, як неформальний необ'єктивний оповідальний характер; використання метафор та стилістично маркованої лексики; відсутність норми на всіх рівнях змісту (тематичний, інформативний, лексичний, синтаксичний); використання слів-тригерів та розпливчастих формулювань; часте вживання прикметників; виразно іронічний або саркастичний тон. Зроблено висновок, що фейкові новини характеризуються більш широким емоційним простором, ніж денотативним.

Ключові слова: постправда, медіадискурс, дезінформація, інформаційна війна, клікбейт.