

Linguistically Manipulative, Disputable, Semantic Nature of the Community Reddit Feed Post

Solomiia Albota

Lviv Polytechnic National University, S. Bandera Str. 12, 79013 Lviv, Ukraine

Abstract

The phenomenon of online community has been deployed in the paper. Its structure, functions and linguistic peculiarities have been regarded. The notion of linguistic manipulation with its classification and general principles has been outlined. The Reddit social virtual platform has served to be a research material. Manipulative nature, disputable features as well as semantic analysis of textual fragments of the Coronavirus Comment discussion section have been considered. Different social network rubrics have been taken into consideration. The Reddit disputable opinion rubric posts and Coronavirus Community comment section were under comparative analysis. The Reddit Coronavirus Community posts have been allocated. They were linguistically marked both implicitly and explicitly for further textual semantic analysis, which was conducted as manually as in a computerized way. The latter provided statistical base for the substantiated linguistic comments to the linguistic and interpretative analysis of the Coronavirus Comment section. The disputable situation analysis considering its conflict solving modes has been applied. The statistical comparison of fragment semantic analysis findings of one of the earlier papers regarding Reddit Coronavirus news comments and this of Reddit Coronavirus Community post has been provided and interpreted.

Keywords¹

Reddit social virtual platform, Coronavirus online community, discussion chart, manipulative nature, disputable situation, linguistic markers, textual semantic analysis, implicitness and explicitness, comparison and contrast

1. Introduction

What is of today's digital world significance – to raise awareness concerning precise, updated data as well as to operate it appropriately up to personal demands and requirements by linguistic and psychological techniques. What became at ease nowadays underlying beneficial consequences is to impose manipulative impact on masses. Indirect implicit influential techniques are considered to be imposed on users within social networks. As any human can be managed on a conscious level of existence, the first and foremost influence he/she is subjected to – linguistic manipulation, which we may regard implicitly or explicitly. Linguistic manipulation is regarded as language peculiarities and set of rules to use it aiming at implicitly affect the social community user with a proper accent. Manipulation is realized when interlocutor is not able to recognize the hidden speaker's intentions in terms of real time conversation. This is mainly vague social and psychological notion, as it triggers detrimental effects on human and society as well. However, it should be kept in mind that such form of desirable lingual determined strategy is supposed to be the most crucial factor in the theories of solving disputes and routine life as well [1,2,3].

COLINS-2021: 5th International Conference on Computational Linguistics and Intelligent Systems, April 22–23, 2021, Kharkiv, Ukraine

EMAIL: Solomiia.M.Albota@lpnu.ua (S. Albota)

ORCID: 0000-0002-0845-5304 (S. Albota)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

When dealing with a notion of social network, one should not neglect the fact that apart from manipulative influences, conversation between users may gain disputable discussion characteristics. If a statement is considered disputable, it will always be manipulative in its essence. Discovering true textual reference to manipulative disputable statements, it is required to overlook spoken background information, but to thoroughly take into consideration each and every detail of the statement uttered and analyze its inner linguistic features. As conversation within social community may be resembled only in a discussion form, it is always common that certain conflict situations remain unresolved. In order to make disputable and manipulative statements flow in their natural way, one must adhere to the tolerant principles of communication process.

2. Related works

Social networks foster lengthy communication process triggering studies based on digital- and software-orientated approaches concerning patterns of human consciousness. In attempt to at least comprehend and slightly uncover the extent of interrelation between mental human inner world and both external and internal language processes, a variety of scientific studies have been considered within the linguistic, psychological, social research areas [1,2,3]. Judging from the personal linguistic features inherent to certain type of speech, one can speculate on interpersonal lingual intentions, determine both negative and positive emotional states of mind, as well as discover general cognitive processes. Since social communities have been discovered in terms of social networking – a basis for instant social cooperation [4], we have to deal with a case of communication exchange within such modern digital environment. When it comes to conversation deployment in social networks, one faces top discussion sections. There, conversation may stream in a form of monologues/dialogues, which will be further extensively commented, thus resulting in polylogues. It depends on users, mainly which type of online community they may generate.

Online communities play crucial role in terms of spreading, sharing/exchanging information. In this way, they are applied to as media for various data challenges. The notion of online community is regarded to be a trigger to establish a reputable social network. It is an open income stream of up-to-date brainstorming, as well as an initiated social media and its technique tool. This notion is of importance as it claims to be interdisciplinary, positioning itself as any form of speech manifestation ability. Notably, there is no single definition of such term as online community. However, online community term may be interpreted as: a certain group of people linked with similar spheres of interests and fields referring to mutual purpose, and those who, for the time being, managed to establish communicative friendly rapport within the environment through Internet technologies [5]. One more definition of online community term regards the fact that this is also a group of people communicating with each other as long as they can chat via web resource. These people are aimed at disseminating received information through different channels of its linguistic interpretation to pursue their own benefits [6]. Such a range of synonyms to the notion of online community as virtual, electronic or even network community can be applied. Any of these synonyms refer to people networking via web services [7, 8]. The only difference that can be seen is that some of these synonymic range may be taken as a lower division of online community. Social groups within this environment interact electronically without facial factors. Such groups communicating by social networking tools are entitled online communities. Their common ground is distinguished as community interaction within web resources, but the difference lies in their structural level. The latter can involve preferences concerning discussion topics, vivid program interface, applied opportunities. If to consider the community network content, one should single out its social networks, discussion letters, web communities [9]. The difference between them depends on the extent of involvement into the internet processes.

Taking into account different social spheres, people usually belong to preferable internet networks, among which virtual game worlds, data exchange, blogging spheres, social networking platforms and social virtuality environment are classified [10]. There are also online communities which can be accessed by everyone. In contrary, there are those which offer their service only for users who have been previously registered. Among other Internet communities, only after obtaining a friendly request can a user become its

member. What is more, types of social networking with textual content, interactive media involvement and visual inclusion are distinguished.

Considering the classification of online communities, the Reddit chatting network (<http://www.reddit.com>) was chosen to become a basis of research. It is referred to social networks as it may involve public among which a vast amount of data is shared. Reddit is marked by a bulk of feeds unfolding rubrics ranging from popular posts (subdivided into hot, new, top around the globe) to Reddit live [11, 12]. In such way, one can choose particular subdivision of feeds based on personal topical preferences. Any user may freely comment a post in discussion chart, share it, download necessary information or just spot the line. This social network may combine both psycholinguistic and social approaches to accumulate patterns in terms of online community conversation. Such an issue of establishing psycholinguistic (manipulative) and social (disputable) features within the Reddit has recently become widely studied [13]. One more essential feature of virtual communities is that discussion on live today's topics is held. When interacting, online community participants along with the information shared make the social media full of senses. Stream of conversation exchanges evolves into discussions, which may be in a form of textual pieces (fragments). Be it sent, discussion chart comment section has its own structure: username, time and date of comment, chosen rubric [14]. Actually, there is a challenge of fragment structure represented in virtual communities [11]: a linear way of textual flow taking into account only time frameworks and branch textual flow showing reference to all previous textual fragment comment provided by other users. Textual fragment comments have as explicit as implicit nature. Explicitly can be called such comments which have extended layout in discussions, branch textual structure, a vast number of references to other users' comments [15, 16, 17]. Implicit comment nature can be revealed by conducting semantic analysis of meaningful statements. Such characteristics may be regarded: constructions being frequently repetitive; keywords characterizing the statement and its synonyms; the number of personal pronouns; linguistic markers determining manipulative and disputable statements, etc.

The main linguistic unit in Reddit social media is fragment comment represented by text. It possesses informative meaning not depending upon context. The comment section in online communities is represented by first comment and the rest of comments. Linguistic influence, which is ignited between these both types of comments, causes multifaceted approaches to its research. The issue was tackled in [18, 20]. The theoretical aspect of speech influence dates back to ancient rhetoric as a science of victory in debate. Social and economic and political development and improvement of the media have led to the need for new means of persuasion with the implementation of psychological and emotional impact on the listener. In modern linguistic studies, language influences the interlocutor, thus regulating personal relationships, mental state and behavior [18]. Speech influence is the object of research not only in linguistics, but also in a number of the related sciences mentioned above. Speech influence in a broad sense is any speech communication in terms of its target conditionality. It is verbal communication, described from the side of one of the communicators, when he/she considers himself/herself as a subject of influence, and his/her interlocutor – as an object [19]. According to O.S. Issers, speech influence implies that the subject regulates the activities of another person, which to some extent is free to choose his/her actions and focuses on his/her needs [20, p. 23].

One of the most effective types of hidden influence is language manipulation. Psychologists interpret manipulation as a type of psychological influence that leads to a hidden motivation of another person to intentions that do not coincide with his/her actual desires [13]. The semantic field of the concept of *manipulation* contains the “negative” intention of the speaker and the hidden (not obvious to the listener) nature of influence. The manipulative functions of discourse generate a hidden layer of linguistic data that is not easy to separate from purely informative content. Verbal manipulation can be seen as a complex, multi-stage, step-by-step procedure (as in the case of advocacy and advertising campaigns). Or else, it can be a single, relatively simple act of influencing the recipient in interpersonal communication. Discourse becomes manipulative not due to the use of specific lexical or grammatical units, but, above all, through the association with the intentions of the speaker, the vague influence of the nature of expression, the conditions of communication (social context). The language has many tools for the implementation of manipulative goals. Linguistic manipulation is distinguished by linguistic markers of different levels, which

help to interpret the intentions of the speaker [21]. Linguistic manipulation is manifested in the verbal address of the speaker, who regulates the behavior of the interlocutor through language, encouraging him/her to enter into a conversation or improve it, supplement if necessary. The speaker can either stimulate an appropriate verbal or nonverbal response, or exert indirect influence to form certain emotions and perception necessary for the speaker. In perspective, the perception is intended to create such behavior that the speaker was targeting, on the part of the listener. By influencing a person, we strive to create his/her behavior in accordance with our needs.

Depending on the sphere of mental activity that participates and dominates in the process of communication, there are rational and emotional types of speech manipulation. In attempt to influence the behavior of the interlocutor, the speaker can influence his/her rational sphere. To do this, he/she uses convincing facts and arguments that affect people's consciousness. The purpose of emotional manipulation is to express the emotions of the speaker and acquire the appropriate emotional response of the listener, which will lead to changes in his/her behavior. Direct emotional manipulation is realized through the initial appeal to the rational side of the listener. Indirect emotional manipulation is performed by creating imagery, error in logical thinking [19]. Depending on the nature of the speaker-listener interaction, direct manipulation involves the fact that the speaker openly states his/her requirements to the listener of the manipulation. Whereas, indirect manipulation is aimed at the environment, not at the listener. Given the speech act, there are intentional or unintentional manipulations. In the case of intentional linguistic manipulation, the subject is directed to a specific result by the object of manipulation. Unintentional language manipulation is involuntary because the subject is not aimed at achieving results from the listener.

According to the type of speech act, manipulation can be the following.

1. Social (socially uninformative speech acts with clichés in the form of greetings, oaths, prayers)
2. Volitional (speech acts, which are accompanied by the will of the speaker in the form of orders, requests, refusals, advice, etc.)
3. Information-evaluation (speech acts that establish public morality, legal interpersonal emotional relations in the form of condemnation, praise, accusations, insults, threats)

The reaction of the addressee is the basis for discrimination of the following types of linguistic manipulation.

1. Evaluative (change of the speaker-listener attitude)
2. Emotional (formation of the general emotional mood)
3. Rational (reconstruction of the categorical structure of personality consciousness, introduction of new categories)

According to the direction of the interlocutor, the manipulation can be focused on the person or society. Personality-oriented language manipulation is aimed at the listener by the speaker, who creates the image of his/her interlocutor to achieve the desired effect. In the case of society-oriented manipulation, the speaker does not create the image of an individual listener, but a generalized image of the group as a whole. Each type of linguistic manipulation can help regulate the activity of the interlocutor and change his/her behavior [19]. The theory of linguistic manipulations involves the differentiation of manipulative and actualized manipulations and productive as well as unproductive manipulations. Such differentiation of means of manipulation within the framework of communication takes the form of a hierarchy, which reflects different levels of communicative skills of language use. Unproductive manipulation occupies the lower part of the hierarchy, while speech actualization is at the top.

In psychology, the term *manipulation* is defined as a type of psychological attachment, which in the case of skillful implementation leads to a hidden provocation of other people's intentions that do not meet his/her real demands and incentives to perform actions necessary for the manipulator [13]. Linguistic manipulation is based on mechanisms that force the listener to perceive verbal messages uncritically and help to create illusions and misconceptions that affect the emotions of the recipient and force him/her to perform actions that are beneficial to the speaker. An unproductive form of manipulative attachment is associated with the desire to secretly influence the consciousness of the interlocutor in order to cause psychological discomfort for him/her. Unproductive manipulation is linguistically aimed at the manifestation of power over the interlocutor, demonstrating his/her imperfection, inferiority, which leads to subordination to the

requirements of the speaker. The purpose of productive manipulation is to win a communicative partner and manipulate his/her behavior through the use of his/her weaknesses. In this case, the initiator of communication becomes a voluntary donor, who positions his/her interlocutor in a situation of social well-being, status advantage.

The simplest means of manipulation are compliments and flattery. Both productive and unproductive manipulations of the addressee's behavior involve the impact on his/her emotional sphere, in contrast to other forms of strengthening the will, which appeal to the rational sphere [19].

The highest form of language interaction and manipulation is communication at the level of cooperative actualization, i.e., the optimal alternative to effective communication. The main behavioral indicator of the actualization agent is respect for the individualism of the interlocutor, fundamental equality and transparency of manipulation methods [21]. Therefore, mastering the actualization of communication is not an easy task.

As a result, manipulative forms predominate in everyday life. One of the most productive types of hidden influence is language manipulation. Techniques of linguistic manipulation are used and will always be used. There will be created new methods of linguistic manipulation, which we simply will not be able to counteract.

3. Methods revealing ambiguous nature of the coronavirus reddit posts

If the conversation exchanges reach their high lever on the verge of conflict situation, one may find ways to manage it effectively. The conflict solving modes mentioned in previous papers [22, 23], though active in oral utterances, may be preferably implicitly applied to textual online community environment. What is worth mentioning is the fact that all disputable statements to be further analyzed are automatically considered to be manipulative, as it was mentioned above (Paragraph 2) that any linguistic influence imposed through statements between users implicitly or explicitly causes hidden or direct type of manipulation. Conducting semantic analysis of the textual fragments in comment discussion charts, particular linguistic markers are considered when marking statements as disputable and manipulative.

The purpose of the paper – linguistically discover cognitive patterns interrelated with implications as well as explications applying linguistic markers within the chosen Reddit comment sections of discussion charts. Afterwards, they will be thoroughly semantically analyzed. Moreover, manual results are to be compared with those conducted automatically applying a software textual analysis program which detects linguistic and psychological markers and patterns, which was previously applied in other articles and proved its analytical efficiency. It is called Linguistic Inquiry and Word Count (LIWC, <https://liwc.wpengine.com/results/>). In earlier papers such Reddit news post rubric as *Coronavirus megathread* was studied [22], as this news firstly loomed on the horizon of news feed. That time, the pandemic process was evolving in its initial stage. Now, it is worth further studies in this news field in order to spot the global virus peculiarities in terms of cognitive shifts and psycholinguistic conversation behavioral patterns. This time, Reddit Coronavirus Community is a study basis. Therefore, there is one more task: to compare the textual semantic analysis of the previous paper [22, 23] with this one taking into account textual interpretative features along with conflict modes and their manipulative embodiment. The disputable modes are applied to both previous and current rubrics along with their comment sections within certain discussion charts. What is more, LIWC software textual analysis initially conducted manually and then automatically was to be applied to this paper and the previous one [23]. Eventual results are considered to be analyzed in a linguistic aspect as well as to be interpretatively contrasted and compared.

In previous articles the attention was paid to the Reddit section of disputable opinions as there one can find a bulk of controversial comments (Fig. 1).

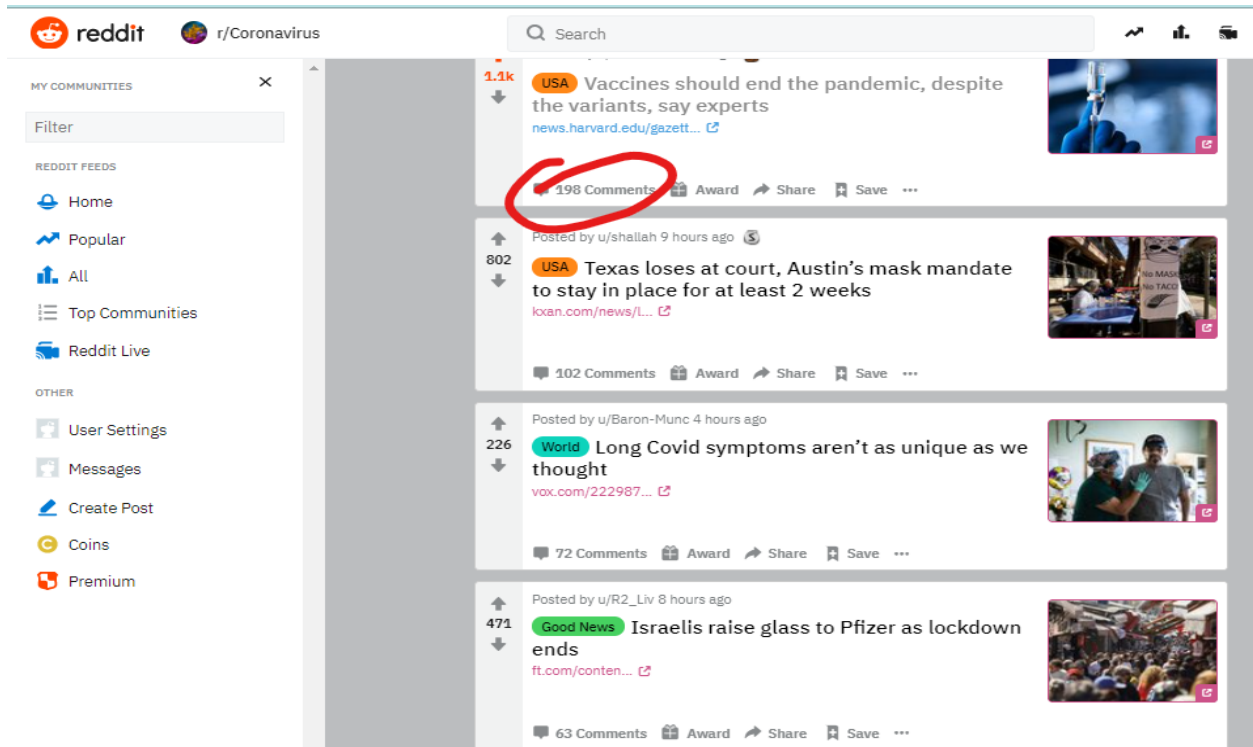


Figure 1: Reddit section of disputable opinions

But this time, the accent was shifted to the Community Reddit feed in order to contrast two converse Reddit structural features. In this Community, subreddit entitled r/Coronavirus (https://www.reddit.com/r/Coronavirus/comments/m44ecb/vaccines_should_end_the_pandemic_despite_the/?sort=controversial), we discovered the most commented rubric with a key word ‘vaccine’ to compare preliminary semantic analysis findings conducted in [22], which were committed to the ‘initial phase’ of Coronavirus, with today’s pandemic ‘final evolving phase’ of Covid-19. Actually, firstly, such entitled rubric was chosen as the most disputable Reddit news feet at the time, but for the time being, it has served to be a reason for further comparison of semantic analysis results. What is more, this Community has an option to be filtered by flair which clarifies the query wanted (Fig. 2). Although we did not specify the flair, our rubric under analysis was chosen automatically to be from USA. The rubric can be sorted by the following comments: best, top, new (suggested), controversial, old, Q&A, out of which we choose controversial to make it easier to trace the disputable nature of comments as well (Fig.3).



Figure 2: The Reddit Community flair-filtered option

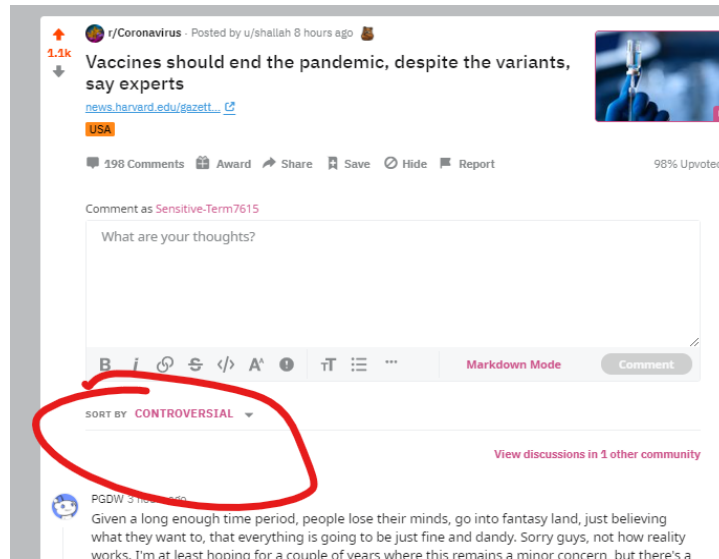


Figure 3: The Reddit disputable rubric comments sorting

Unlike Reddit news section, Community members make a reference to the web source, mark it in order to further comment and expand into lengthy discussion (Fig.4).



Figure 4: Reference to the outward web source

In our case, the article is taken from *The Harvard Gazette*, which is marked by rubric *Health & Medicine*. One more opposite structural characteristic is that discussion charts do not have titles as there is one common Coronavirus Community. This can be considered as disadvantage, since apart from semantic analysis, we have to conduct preliminary linguistic contextual analysis aiming at implicitly and explicitly reveal potential headings for discussion charts. This way, we allocate the most frequent linguistic markers with manipulative and disputable implications according to the following headings.

1. **Mental virus-vaccine effects**
2. **Virus-vaccine ‘geography’**

Such heading allocation was firstly performed manually, using linguistic interpretive analysis, contextual analysis revealing implicit and explicit nature of discussion utterances (predominantly in phrases, sentences and full passages), which was later successfully proved automatically.

The first heading – **mental virus-vaccine effects** – involves social manipulative psycholinguistic material concerning Community member’s speculations on how Coronavirus and vaccination affect society in general. The discussion thread topics range from deep psychological problems, hopes and predictions to primitive typical chats.


The second heading – **virus-vaccine ‘geography’** – covers location-centered manipulative statistics and places of expansion with fearful facts and debates on warning and taking actions. Here, the accent is made on current ongoing epidemic situation.

Currently, by conducting semantic analysis, we allocate semantic groups according to the previous headings, which are grouped in accordance with those introduced by the automatic LIWC tool (Fig.5 and Fig.6).



TRADITIONAL LIWC DIMENSION	YOUR DATA	AVERAGE FOR SOCIAL MEDIA: TWITTER, FACEBOOK, BLOG
I-WORDS (I, ME, MY)	1.5	5.51
SOCIAL WORDS	11.1	9.71
POSITIVE EMOTIONS	3.3	4.57
NEGATIVE EMOTIONS	3.3	2.10
COGNITIVE PROCESSES	18.2	10.77

Figure 5: Computerized semantic analysis of the Reddit Coronavirus Community comment section in the first heading



TRADITIONAL LIWC DIMENSION	YOUR DATA	AVERAGE FOR SOCIAL MEDIA: TWITTER, FACEBOOK, BLOG
I-WORDS (I, ME, MY)	1.4	5.51
SOCIAL WORDS	8.9	9.71
POSITIVE EMOTIONS	2.9	4.57
NEGATIVE EMOTIONS	1.6	2.10
COGNITIVE PROCESSES	13.7	10.77

Figure 6: Computerized semantic analysis of the Reddit Coronavirus Community comment section in the second heading

We apply conflict solving modes previously used in [22] to the Coronavirus Community comments. The first discussion heading chart is shown below. It is followed by linguistic markers (in bold, italics) implying.

1. **Negative emotions:** *people lose their minds* (mental problems+**personal factor**), (accommodating, avoiding); *go into fantasy land* (mental problems+**personal factor**), (avoiding); *just believing what they want to* (mental problems+**personal factor**), (avoiding); *this remains a minor concern* (mental problems), (collaborating); *If we can't get it wiped out, part of the world will remain a petri dish for better versions of the virus* (negative prediction), (collaborating), *unless you're dumb enough to throw a massive part* (mental problems+**personal factor**+**cognitive processes**), (competing); *Wow not having anyone in your yard is crazy* (**negative prediction**), (collaborating); *It's been hard*

mate, trust me (**negative prediction**), (avoiding); *someone dealing with a severely mentally ill family member* (mental problems+**personal factor**), (avoiding); *The mental health support is so bad* (negative prediction), (avoiding)

2. **Positive emotions:** *everything is going to be just fine and dandy* (promising hopes), (collaborating); *Try reading the more positive news articles* (advice+**cognitive process**), (collaborating); *why can't we open everything, when will we open..it might change if data proves* (promising hopes), (collaborating); *they just keep saying he seems to be fine* (promising hopes), (collaborating); *Its totally possible this thing ends via vaccines* (promising hopes), (collaborating); *We're going to go back to living our lives* (promising hopes), (collaborating); *we just have to accept the level of risk and move on* (promising hopes), (collaborating); *We take appropriate precautions and move on with our lives* (promising hopes), (collaborating); *Paradise is on the horizon* (promising hopes), (collaborating); *but yes it will help a lot in suppressing the virus out of an active pandemic state* (promising hopes), (collaborating); *begin to come up with a long term recurring vaccination plan* (promising hopes), (collaborating); *Well I'm pretty sure that's the plan* (promising hopes+**personal factor**), (collaborating); *Forget the vaccine and COVID* (promising hopes), (collaborating); *I'm still hopeful that the vaccine efficacy is enough to end..* (promising hopes+**personal factor**), (collaborating); *we will hopefully adapt to coexist with COVID* (promising hopes), (collaborating)

3. **Cognitive processes:** *Sorry guys, not how reality works* (explicitness of instability+**negative emotions**), (avoiding), *but there's a real risk that this becomes perpetual* (implicitness of instability+**negative emotions**), (avoiding), *Mental health needs to be taken more seriously* (explicitness of instability+**negative emotions**), (collaborating); *Even the most cautious people .. having people over outdoors pretty much the whole time* (explicitness of instability+**negative emotions**), (collaborating); *Many if not most people ignore that rule* (**works** (explicitness of instability+**negative emotions**), (avoiding); *There are many things that "could" kill you tomorrow* (explicitness of inevitability+**negative emotions**), (accommodating)

Manually, more than 100 textual fragments of the first discussion chart were thoroughly analyzed, but it was given partially above to generally illustrate the process of semantic textual analysis. Such features as reference to personal factor and social words are discovered automatically but manually there was no necessity to allocate them separately as they were considered while analyzing (e.g., marked with pluses). This way, we prove with an automatic program report that there is an equal number of positive emotions and negative ones in the comments. It can be substantiated that such Reddit rubric is neither positive nor negative, we have a case of mere mental processes aiming at bridging the gap between long-term pandemic stagnation and finding ways to eradicate that. The increase in social words factor may be substantiated by frequent semantics of personal mental problems marked linguistically.

To compare the average percentage results of the first heading concerning the Reddit comments semantic analysis with our data (Fig.5), one should note no particular frequency of any semantic field subdivision (social words, positive emotions, etc.) – the percentage figures are even lower than the average amount. Such instability in textual comment fragments prove general insufficient attention paid to such Reddit rubric, what cannot be said about cognitive processes. They are ranked higher than the average percentage, what is proved both manually and automatically. The Reddit Coronavirus rubric debates have been actively considered involving factual base of statements following personal deductions, which are proved by singling out linguistic markers denoting them.

The second discussion heading chart is shown below. It is followed by linguistic markers (in bold, italics) implying.

1. **Negative emotions:** *Yah but that's a personal choice so they can get sick* (fear+**personal factor**), (accommodating, avoiding); *No shit, they're wanting to keep us locked down for as long as they want...* (fear+**personal factor**), (avoiding); *Enforcement of this is obviously pretty low though* (fear), (avoiding); *Tired of the "maybe small gatherings by July" bullshit* (fear+**personal factor**), (avoiding)

2. **Positive emotions:** *that will likely only be true for places with easy access to super high efficacy vaccines* (promising hopes), (collaborating); *and has the ease-of-deployment advantage* (promising hopes), (collaborating); *and we're smashing targets for vaccinations every week* (promising hopes),

(collaborating); *The UK is doing very well with their vaccination pace* (promising hopes), (collaborating); *There are well over 4,000 variants already sequenced* (promising hopes), (collaborating); *I'm happy to live in one that is mostly open now* (promising hopes), (collaborating); *I'm glad it'll end* (promising hopes), (collaborating); *Personally, I'm looking ahead to near normal conditions* (promising hopes+personal factor), (collaborating); *perfectly expected that we're figuring out that the vaccines are working well* (promising hopes), (collaborating); *Seems perfectly fair* (promising hopes), (collaborating)

3. **Cognitive processes:** *if ...controlled the virus, we'd have an accurate guess* (explicitness of action+**positive emotions**), (collaborating); *Am I the only one seeing a bit issue in this statement?* (explicitness of action+**positive emotions**), (collaborating); *I really think mask mandates should remain in place in certain industries* (explicitness of action+**positive emotions**), (collaborating)

Manually, more than 110 textual fragments of the first discussion chart were thoroughly analyzed, but it was given partially above to generally illustrate the process of semantic textual analysis. Such features as reference to personal factor and social words are discovered automatically but manually there was no necessity to allocate them separately as they were considered while analyzing (e.g., marked with pluses). Here, we prove with a computerized report that positive emotions prevail in the comments, as applied linguistic markers helped to determine semantic coloring of hope for normal life conditions involving vaccination.

In attempt to contrast the balanced ratio outcomes of the first heading concerning the Reddit comments semantic analysis with our data (Fig.6), one should note no particular frequency of any semantic field subdivision (social words, positive emotions, etc.) – the percentage figures are even lower than the average amount (Table 1). Such instability in textual comment fragments prove general insufficient attention paid to such Reddit rubric, what cannot be said about cognitive processes. They are ranked higher than the average percentage, what is proved both manually and automatically. The Reddit Coronavirus rubric debates have been actively considered involving factual base of statements following personal deductions, which are proved by singling out linguistic markers which helped to denote them.

Table 1

Table of statistical figures concerning the Reddit Coronavirus Community comment sections in both headings

Comment section	I-words (I, me, my)	Positive emotions	Negative emotions	Cognitive processes
1	1.5	3.3	3.3	18.2
2	1.4	2.9	1.6	13.7

To sum up, the two allocated headings – mental virus-vaccine effects and virus-vaccine ‘geography’ – have manipulative, disputable nature as it was discovered linguistically, mainly semantically and more explicitly than implicitly. All semantic fields singled out (mental problems, negative prediction, promising hopes for future, advices, instability situation, inevitable processes, fear and taking action) are being manipulated in a disputable way between users’ statements and mass media as well. Specific personal accents were linguistically marked and spotted manipulatively in the first heading, as people are bombarded from each corner with the same informative volume, thus affecting personal state of mind with future health consequences. The semantic group of positive emotions prevails as we face hopes for ‘healthier’ life in future than positive virus-vaccination prediction. More negativity is expressed linguistically through personal health experience than manipulative personal fearful states. Concerning cognitivism, first heading semantics is linguistically marked by warnings, threats, consciousness manipulations rather than taking action plans, manipulative rhetoric questions in the second heading comment section.

4. Contrasting both reddit disputable situation results

In earlier papers [22, 23], studies were conducted concerning the disputable Reddit comment nature and comment discussion charts were under deep analysis on textual semantic analysis with an application of linguistic markers to drastic conflict resolution. One comment section we have to compare with the one under analysis (Table 3) is *Coronavirus megathread* rubric discussion chart (Table 2).

Table 2

Table of statistical figures concerning disputable modes of the Reddit news conversations about Coronavirus

Comment section	Compromising	Accommodating	Competing	Avoiding	Collaborating
1	0	1	2	2	2
2	0	2	3	1	1
3	0	1	1	1	4
4	0	2	3	0	2

Table 3

Table of statistical figures concerning disputable modes of the Reddit Coronavirus Community comments in both headings

Comment section	Compromising	Accommodating	Competing	Avoiding	Collaborating
1	0	2	1	9	20
2	0	1	0	4	13

First of all, a question arises in terms of possibility to compare 4 post comment sections and current 2 comment sections. The answer is that statistically the amount of textual fragments accounts for their similar number not less than 100 discussion charts. Moreover, one should not forget that Table 2 represents post comment sections on the basis of the Reddit *Unpopular opinion* rubric unlike Table 3 shows the Reddit Coronavirus Community comments, which do not have such an automatic function to group comments by topics (headings). That is why, we have to interpret each discussion chart comment to make it divided into 2 comment section. There was no reason to classify the comments into more sections as all cognitive processes in this Reddit Community involve 2 main spheres: virus-vaccine mental effects and virus-vaccine ongoing situation.

Let us analyze the comment section of the current Reddit Coronavirus Community discussion chart. As the topic of Coronavirus has a negative implication in its essence, all statements referring to it render disputable, mainly manipulative features. Such conflict resolution strategy as compromising was not applied to none of the previously allocated textual fragments as there was nothing to negotiate, users mainly firmly state their opinions without seeking support or mutual common ground – that was rather a topic for inner ‘cry’. Accommodating technique was inherent to both first and second comment sections. Both techniques of accommodating are implicitly revealed as they, in most of their cases, denote accepting of the real facts, following the current state of mind involving personal factors. None of the analyzed discussion charts had been applied to competing conflict resolution strategy but one manipulative example illustrating challenge to get together in pandemic. The avoiding strategy is inherent to negative emotions, desire to sidestep health problem issues. The first comment section is more frequent in inner cognitive processes to quit, escape reality, inability to stand current situation than the second one, which is uncovered by certain linguistic markers denoting virus restrictions and rude behavioral patterns. Analyzing content input and identifying main current concerns is inherent to collaborating conflict resolution strategy, frequency of the first of which is higher than that of the second one, as internal human desires strive for

mutual efforts to persist the pandemic mood are fairly on the verge of taking actions rather than expressing promising hopes, which are explicitly revealed in the second Community comment section.

Now, to contrast and compare our current results concerning application of conflict solving modes with the previous one (Table 2), we may firmly agree with the fact that no compromising resolution was applied to the allocated Coronavirus rubric textual fragments. What concerns accommodating resolution strategy, the same frequency figures were considered, which proves the similarity of acceptance of the ongoing process concerning current pandemic situation. The main manipulative feature in this term of textual fragments being applied to accommodating strategy is mere quitting one's own desires to persist virus obsession. The strategy of competing is much higher, as dramatically assertive linguistic markers were revealed denoting given choice to be alive or die, to get infected on purpose, etc. Though the same negative virus semantics dominated in the previous rubric, the frequency of application of avoiding strategy is comparatively low. It may be substantiated by the fact that no reality escape was predicted, all issues were shifted to whether collaborative or competing strategy modes. Collaborating conflict resolution is pretty low in comparing frequencies, as there was no strong necessity to support mutual inability to take actions or maintain conversation rudeness.

The most distinctive feature between those two general comment sections is the fact that statements singled out in the previous paper on Coronavirus rubric were represented in a form of words, phrases involving exclamations, linking words, rude vocabulary denoting speculations on probable virus situation scenarios. As users were not so profoundly aware of the pandemic as they are nowadays, their form of discussion was abrupt, rhetoric, required lots of contradictions. Current textual semantic analysis proved that statements in Coronavirus Community comment sections resemble a form of sentences, which we had to contract in order to vividly demonstrate those essential marked parts. This way, sentences were more substantiated, as users seemed to operate factual base of the situation for the time being, thus manipulating users' consciousness in a way to provoke disputes on issues. But such drastic disputes as in the previous analysis were not spotted, as facts and knowledge base concerning virus-vaccination initial and ongoing phases hindered conflict situations and fostered collaboration.

5. Discussion on the social platform comment analysis considering semantic and statistical aspects

This paragraph is based on the preliminary results discovered above (Paragraph 3, Table 1) in order to compare and contrast results from the previous paper [22], (Table 4) taking into account semantic groups in both comment sections.

The first semantic group that was revealed is reference to personal factor. Similar to our current findings, no personal pronoun peculiarities were revealed. Even if statements referring to personal factors in both findings being involved into such semantic groups as negative and positive emotions, which was revealed manually, no particular manifestation was proved by software application (0 cases).

The second semantic group is entitled negative emotions, which in previous research had 19 cases in total of each comment section, in current analysis – 23 cases, which proves more awareness of the terrifying consequences of virus we face nowadays. These results were proved manually and by program tool.

The third semantic group is called positive emotions, which previously fixed 12 cases and currently we have 24 of them. It is substantiated by interpreting the fact that today users hope for better situation as vaccine appeared and more restrictions are being unlocked. These results are more reliable using manual analysis as automatic tool, in our case, proved equal amount of positive and negative emotions in the second rubric heading, which seemed to be a misleading way of proof.

The fourth semantic group is referred to cognitive processes, which in early paper had 9 cases which is equal to our analysis with the same number of 9 cases and is both manually and automatically proved.

Eventually, let us contrast and compare frequency results established by LWIC software tool helpful in conducting textual semantic analysis of the Reddit both rubric and community comment sections within

discussion charts. Table 5 is represented by total amount of frequency, where the first discussion chart refers to previous rubric results, and the second – current Community comment section.

Table 4

Table of statistical figures concerning the Reddit comments within the Coronavirus discussions

Comment section	I-words (I, me, my)	Positive emotions	Negative emotions	Cognitive processes
1	3.3	4.7	6.0	19.3
2	1.7	2.3	5.1	10.2
3	3.4	7.8	3.4	18.1
4	3.1	2.5	1.3	11.9

Table 5

Table of statistical figures concerning the Reddit Coronavirus Community comments in both headings

Discussion charts	I-words (I, me, my)	Positive emotions	Negative emotions	Cognitive processes
1	1.8	4.3	3.7	15.9
2	0.9	8.7	11.8	99.0

Principally, the study is considered reliable if the software tool findings and manual textual semantic analysis have a common ground and are not distorted. Nevertheless, such language inclination and linguistic competency allows us to presume that manual findings tend to be more accurate. Involving manipulative nature, disputable features of each statement allocated and thoroughly linguistically analyzed, we can state that automatic tool proved manual analysis results precisely.

I-words prevailed in the first discussion charts as short and abrupt statements are inherent to title-rubric comment section, whereas within communities, opinion-centered approach is concerned. Positive emotions dominated in the second discussion chart as both linguistically explicit and implicit manipulations of promising ‘bright’ future in terms of mental Coronavirus health effects were greatly detected. Negative emotions prevail in the second discussion chart being marked linguistically as fearful efforts to overcome virus effects and its further mental problems. Unlike cognitive processes in the first discussion chart, the second ones were substantiated mainly by the implicitly revealed instable health modes and explicitness of taking actions to remain ‘unlocked’ under usual life conditions.

6. Conclusion

Ultimately, the linguistic investigation involved such aspects as general manipulative nature of textual fragments and their disputable features. Contextual, linguistic and interpretative as well as textual semantics analysis were applied to the Reddit Coronavirus Community comment section within the discussion charts. Online communities were defined as a variety of social platforms for data dissemination, as well as conducting debates followed by commenting and different types of posting opportunities. The notion of linguistic manipulation, its features, types and general principles were considered.

This time, Reddit discussion comment sections were chosen not from random *Unpopular opinion* rubric but from Coronavirus Community rubric. Such choice was substantiated by required contrasting and comparison of the previous findings on initial phase of Coronavirus rubric discussion charts with ongoing phase of virus-vaccine effects and mental health consequences. Concerning textual semantic analysis, semantic groups were singled out on the basis of linguistic comments, which were marked, analyzed both implicitly and explicitly and proved as manually as statistically. For the latter purpose, such software tool as

LWIC revealing linguistic markers and psycholinguistic patterns' encoding and decoding was applied. After each paragraph justified linguistic comments were given.

The Reddit disputable situation analysis which implies solving disputable modes, represented in appropriate Tables, was performed. In addition, Tables of frequency of computerized fragment analysis of the Reddit Coronavirus comments were provided. The statistic comparison of semantic analysis of both Reddit Coronavirus posts was performed.

In psycholinguistic perspective, social models of communicative behavior may be a ground for discovering peculiarities of online communication affected by pandemic.

7. References

- [1] M. Carroll, C. Stutterheim, R. Nüse, The language and thought debate: A psycholinguistic approach, in: Th. Pechmann (Eds.), *Multidisciplinary Approaches to Language Production*, 2001, pp. 183–218. doi:10.1515/9783110894028.183.
- [2] Ch. Reuter, S. Stieglitz, M. Imran, Social media in conflicts and crises, *Behaviour & Information Technology* 39:3 (2020) 241–251. doi:10.1080/0144929X.2019.1629025.
- [3] R. C. Farrugia, Facebook and Relationships: A Study of How Social Media Use is Affecting Long-Term Relationships, Master's thesis, Rochester Institute of Technology, 2013.
- [4] T. Althoff, P. Jindal, J. Leskovec, Online Actions with Offline Impact: How Online Social Networks Influence Online and Offline User Behavior, in: *Proceedings of the 10th. ACM International Conference on Web Search and Data Mining, WSDM '17*, Association for Computing Machinery, New York, NY, 2017, pp. 537–546. doi:10.1145/3018661.3018672.
- [5] G.J. Qi, C. C. Aggarwal, T. S. Huang, Online Community Detection in Social Sensing, in: *Proceedings of the 6th. ACM international conference on Web search and data mining, WSDM '13*, 2013, pp. 617–626. doi:10.1145/2433396.2433474.
- [6] A. J. Kim, *Community building on the Web: secret strategies for successful online community*, 2006.
- [7] W. Zhang, Research on the English Online Learning Community Based on SNS, in: *Proceedings of International Conference on Education, Management and Computing Technology*, 2015, pp. 1841–1844. doi: 10.2991/ICEMCT-15.2015.385.
- [8] A. Peleshchyshyn, O. Tymovchak-Maksymets, Analysis of the communicative interaction of participants of community web-forums, *Eastern-European Journal of Enterprise Technologies*, (6 (8 (48)) (2013) 44–49.
- [9] A. Peleshchyshyn, O. Markovets, V. Vus, S. Albota, Identifying specific roles of users of social networks and their influence methods. in: *Proceedings of the 13th. International Scientific and Technical Conference on Computer Sciences and Information Technologies*, 2018, pp. 39-42. doi: 10.1109/STC-CSIT.2018.8526635.
- [10] A. Kaplan, M. Haenlein, Social media: Back to the roots and back to the future, *Journal of Systems and Information Technology* 14 (2) (2012) 101-104. doi:10.1108/13287261211232126.
- [11] G. Haralabopoulos, I. Anagnostopoulos, Sh. Zeadally, Lifespan and propagation of information in Online Social Networks: A case study based on Reddit, *Journal of Network and Computer Applications* 56 (2015) 88–100. doi:10.1016/j.jnca.2015.06.006.
- [12] H. Mensah, Lu Xiao, S. Soundarajan, Characterizing the Evolution of Communities on Reddit, in *International Conference on Social Media and Society, SMSociety'20*, Association for Computing Machinery, New York, NY, 2020, pp. 58–64. doi:10.1145/3400806.3400814.
- [13] S. Pandrekar, X. Chen, G. Gopalkrishna, Social Media Based Analysis of Opioid Epidemic Using Reddit, in: *Annual Symposium proceedings, AMIA Symposium*, 2018, pp. 867–876. pmid:30815129; pmcid:PMC6371364.
- [14] A. Y. Wang, M. H. Newlin, T. L. Tucker, A Discourse Analysis of Online Classroom Chats: Predictors of Cyber-Student Performance, *Teaching of Psychology*, 28:3 (2001) 222–226. doi:10.1207/S15328023TOP2803_09.

- [15] T. Bogers, R. Nordenhoff Wernersen, How 'Social' are Social News Sites? Exploring the Motivations for Using Reddit.com, in: Proceedings of the iConference, iSchools, 2014, pp. 329–344.
- [16] V. Vysotska, A. Sachenko, V. Lytvyn, Y. Burov, Information resources processing using linguistic analysis of textual content, in: Proceedings of the 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS, Bucharest, 2017, pp. 573–578. doi:10.1109/IDAACS.2017.8095038.
- [17] I. Khomytska, V. Teslyuk, A. Holovatyy, O. Morushko, Development of methods, models, and means for the author attribution of a text, Eastern-European Journal of Enterprise Technologies, 3 (2 (93)) (2018) 41–46.
- [18] A. V. Moiseeva, A. S. Titlova, Manipulative influence of microblog text (as exemplified by the comments on twitter), Vestnik Bashkirskogo Universiteta 24 (4) (2019) 996–1002. doi:10.33184/bulletin-bsu-2019.4.39.
- [19] E. L. Dotsenko, Psychology manipulation: phenomena, mechanisms and protection, 3rd. ed., M., 2003.
- [20] O. S. Issers, Speech impact in terms of cognitive categories, Vestnik Omskogo un-ta 1 (1999) 74–79.
- [21] J. Thomas, Meaning in Interaction: An Introduction to Pragmatics, 1st. ed., 1995.
- [22] S. Albota, Resolving conflict situations in reddit community driven discussion platform, in: Proceedings of the 4th. International conference on computational linguistics and intelligent systems, COLINS 2020, CEUR Workshop Proceedings, 2604, 2020, pp. 215–226.
- [23] S. Albota, Linguistic and Psychological Features of the Reddit News Post, in: Proceedings of the 15th IEEE International Conference on Computer Sciences and Information Technologies, CSIT, Zbarazh, 2020, pp. 295–299. doi: 10.1109/CSIT49958.2020.9321991.