Representation of Internet privacy issues in news reports: Critical Discourse Analysis

Darya A. Aripova
Irkutsk National Research Technical University,
Institute of Linguistics and Intercultural communication,
Department of Foreign Languages for Technical Specialties №1
Irkutsk, Russia
aripovad@mail.ru
https://orcid.org/0000-0002-4515-4834

Irina S. Bashmakova
Irkutsk National Research Technical University,
Institute of Linguistics and Intercultural communication,
Department of Foreign Languages for Technical Specialties №1
Irkutsk, Russia
bis61@mail.ru
https://orcid.org/0000-0003-0194-5519

Abstract - This paper aims at exploring morphological, syntactic and semantic features through which social actors, implicated in Internet privacy issues discourse, in news reports are represented. Van Leeuwen’s theoretical-methodological tools have been used in the study to reveal the basic tendencies that conceptualize social interaction between participants involved in the process of privacy violation. Groups of dominant social actors and their roles in digital surveillance practices have been specified by means of the three sets of categories of inclusion/exclusion, activation/passivation, personalization/impersonalization. The obtained results can be used for enhancing readers’ critical thinking abilities in order to understand and create discourse in the media.

Keywords - Critical discourse analysis; privacy issues; social actors; recontextualization; exclusion; backgrounding

I. INTRODUCTION

The Internet environment has a strong and uncontrolled impact on all spheres of individual and social consciousness. With the increase in the number of users of the global network, it is becoming not only a media source, but also a space that contains a huge amount of confidential information. The information can be efficiently collected, processed and analyzed without consumer consent [1]. Due to this fact, the prevalence and growth of the global and yet largely unregulated Internet challenges the personal privacy issues. Information privacy is defined as an aspect of information technology and data security that deals with the controlled usage, circulation and tracking of personal information [2]. Concerns for information privacy comprise quantitative methodology for determination of users’ tracked visits to web sites, e-mail addresses and other personal information marketing and captures, improper access, credit card thefts, unauthorized secondary use of data in online environments in a multidimensional scale [3], [4]. News is considered as institutionally, or man-made social practices revealing stances of its producers. Given the large number of newspapers and magazines in the world, people’s conceptions of social and cultural phenomena can be consciously or unconsciously shaped and reshaped by media controllers. On account of the political and sociological ramifications of news the reality is interpreted and reported in various aspects. News reports reflect the attitude, interests and values of individuals or the institutions they produce and have the evident meaning-making function [5].

Critical Discourse Analysis (CDA) is a branch of linguistics that simultaneously suggests three dimensions of analysis: text, discursive and social practices. This approach is based upon the assumption that there is considerable congruity between language and other elements of social life it represents [6], [7]. Thus, the order of a discourse is interconnected with network and structure of social practices [8]. Fairclough emphasizes the significance of correlation between discourse and society that establishes social representations, features, beliefs and knowledge [9]. Van Leeuwen acknowledges that all speakers and writers operate from specific discursive practices originating in special interests and aims which involve inclusions and exclusions [10]. Van Dijk contends that CDA is able to make a significant contribution to social situation and identities if it takes into account the role of language, language use, participants and their properties, discourse and communicative events [11]. A text is a depository of sociocultural practices and their effects [12], as well as a material manifestation of a discourse [13]; hence texts constitute the data and objects of analysis for discourse analytical studies. The purpose of this article is to investigate representations of Internet privacy issues by means of principles and models of CDA outlined by the above-mentioned theorists and make an attempt to interpret, explain and identify relations of social actors in news reports in question.

II. MATERIALS AND METHODS

This study is based on linguistically oriented conceptual framework proposed by Theo van Leeuwen. In his seminal book on the representation of social actors, van Leeuwen capitalizes on the assumption that all discourses with their key elements can recontextualize social practices. Specifically, social actors, their roles, identities, behavior, actions, reactions, performance styles, settings and timings may be recontextualized. This process presupposes retrieval and transformation of certain discursive elements, as well as focus on purposes, legitimation for, or critique of some actions. The author proposes the mapping of the linguistic elements of a text and the capacity to make clear the actions and their actors. During transformations of the social practices elements substitution of real social actors with semiotic elements occurs. Van Leeuwen categorizes the representation of social actors as a set of linguistic units related to agency which can function to include or exclude people or groups. Thus, the scholar has distinguished two basic types of representation: inclusion and exclusion.
Direct nomination, specification and individualization of participants materialize inclusion representation. Exclusion representation is defined in case when there is suppression with no reference to a social actor in the text, or grounding (concealment) of a social actor. Participant’s placement in a second plan is termed as grounding, partially excluded social actors meanwhile may not be mentioned in relation to a certain activity, but they are referred to in other parts of the text and can be retrieved. Nominalizations, process nouns, generalizations, abstractions, nonfinite and attributive clauses are notably attributed to complete or partial exclusion [14]. In the present paper Theo van Leeuwen’s framework of the representation of social actors, as a powerful tool of examination in CDA, is adopted and applied to analyzing data withdrawn from news reports.

Particularly, the study tries to get the answers to the following questions:

1. What are the representations of fraudulent Internet activities in the media discourse?
2. How is public opinion on Internet privacy concerns implicated in the news reports?

The materials for this analysis were taken from online papers: The Guardian [15], [16], [17], [18], [19], International Business Times [20], The Verge [21], [22], [23], [24] (2018-2019), by means of a continuous sampling method, all dealing with Internet privacy issues and personal data regulations.

### III. RESULTS AND DISCUSSION

Let us analyze the following contexts in order to exemplify the presence of recontextualized participants, their roles, actions and reactions in question.

1) A whistleblower came forward in March to reveal that Cambridge Analytica improperly harvested the personal Facebook data of 50 million people in order to profile and target users for political advertisements. After five days of silence, Mark Zuckerberg apologised for his company’s “mistakes”.

2) A bug in June caused a glitch that publicly published the posts of 14 million users that were intended to be private. Erin Egan, Facebook’s chief privacy officer, apologised for the “mistake” and said the company had since fixed the glitch.

3) Hackers were able to access and steal personal information in nearly half of the 30m accounts affected in a September security breach. The company initially believed that 50 million users were affected in an attack that gave the hackers control of accounts.

“50 million people” and “14 million users” are assimilated and subjected to the actions of “targeting” their “personal data”, “publishing private posts” and “stealing personal information”. By enumerating the approximate number of people (“50 million” and “14 million”) the author of the article emphasizes the quantitative impact of the process. A British political consulting firm and political organizations are associated to form the group opposed to the interests of “users” and benefited in relation to “harvesting the personal Facebook data.” “A bug” and resulting from it “glitch” manifest instrumentalization that backgrounded “hackers” in extract 3. An abstraction and generalization – “a whistleblower” represents an unknown social actor by means of an action “to reveal” an unlawful or immoral activity. Mark Zuckerberg’s and Erin Egan’s actions are realized as 1) an objectivated action (“five days of silence” – serving as a process noun substituting for “keeping silence”), 2) verbal action (“said”) and 3) remedial semiotic reactions (“apologised”). The objectivation of Facebook chairman’s action serves to add legitimation with the negative evaluation. The rendition of “mistakes” by means of quotation marks reveals the true meaning of this notion – that is a means-oriented substantive to “steal personal information”.

A. Inclusion/exclusion pattern

For the purposes of analysis we shall determine “users” as those social actors who exploit Internet and unify “us” and “you” as well. We shall refer to Facebook, Google, Microsoft, Amazon and Apple as “company management”, to the various political consulting firms, surveillance organizations, advertising companies and hackers as “private data seekers”, and to the representatives of the Federal Trade Commission (FTC) and other privacy regulators as “privacy organizations”.

If to consider more closely, recontextualization of included/excluded actors within the examined data provides noteworthy results. As table I indicates, the most frequently backgrounded or suppressed social actors are “company management”. Whereas the Facebook leaders are concealed in terms of linguistic manifestation, it is only them who are entitled to “admit” that the company conducted experiments on its users:

4) Facebook admitted it ran experiments on hundreds of thousands of users to test how emotions spread on the site.

“Users”, who are not mentioned in the sentences below, negatively benefit from the remedial actions of Facebook managers who “apologized for mistakes” of the company. Consequently, suppression is realized by means of deemphasizing social actors to whom the Facebook executives apologized:

5) After five days of silence, Mark Zuckerberg apologized for his company’s “mistakes”.

6) Erin Egan, Facebook’s chief privacy officer, apologized for the “mistake” and said the company had since fixed the glitch.

7) Most recently, it was discovered that Facebook users could be found if someone agreed to the two-factor authentication system and used their phone number.

Some agency or group of people could possibly state privacy violation of Facebook users but there is no reference to these individuals or institution in question, therefore “privacy organization” role is suppressed in the discourse. Exclusion is realized through passive agent deletion – “it was discovered”.
The mostly emphasized actors are “private data seekers” coming second in percentage terms only to “privacy organizations”.

An objectivated action realized by nominalizations “re-election campaign” functions as “personal data acquirer” and the subject of a clause. “Private data seekers” remain linguistically opaque in this case:

8) We knew that in 2012 the re-election campaign of Barack Obama had built a voter contact system using Facebook and had acquired personal data on millions of American voters.

B. Role allocation pattern

Table II shows how social roles are allocated in the online press covering Internet privacy concerns. The activation/passivation pattern confirms that users are the most passivated social actors among the defined participants due to the fact that they experience goal-oriented actions and provide information to establishments and people deriving benefits from it. While organizations searching for private data (“third-party developers”) are beneficialized with regard to having “the ability to access users’ photos,” “users” are subjected to the material transaction of the instrumentalized nomination “bug”:

9) The bug allowed third-party developers the ability to access users’ photos, including those that had been uploaded to Facebook’s servers but not publicly shared on any of its services.

Let us consider one more example of users’ subjection realized by their participation in a forcibly interactive process of “sharing” private data with unknown information thieves:

10) Modern online devices and services are built and funded through the information we share with them. Apps that help you get around are probably reporting on where you go, and devices that turn your words into actions are probably reporting on what you say. It became clear in 2018 that data reveals a lot about us, in ways that create some really big concerns.

In this context the process of getting information from the users (nominated as “we” and “you” in collective meaning) is presented as goal-oriented. Security and intelligence agencies are referred to by means of the words “apps” and “devices”, the former ones are not directly included in the text, but their actions are apparently explicit.

The actions of the objectivated actors appropriating personal information represent semiotic transactions through prepositional verbs “report on” and “turn one’s words into” in which users are passivated and subjected to the state of personal data disclosure (“data reveals a lot about us”), whilst the surveillance agencies capitalize on it.

| TABLE I. INCLUSION/EXCLUSION IN THE NEWS REPORTS REGARDING PRIVACY |
|------------------------|-----------------|-----------------|-----------------|
|                        | Included         | Backgrounded     | Suppressed      |
| "Users" (N=36)         | 58.33%           | 27.78%           | 13.89%          |
| (N=21)                 | (N=10)           | (N=5)            |                 |
| "Company management" (N=51) | 39.22%         | 56.86%           | 3.92%           |
| (N=20)                 | (N=29)           | (N=2)            |                 |
| "Private data seekers" (N=47) | 65.96%         | 27.66 %          | 6.38%           |
| (N=31)                 | (N=13)           | (N=3)            |                 |
| "Privacy organizations" (N=42) | 80.95%         | 16.67%           | 2.38%           |
| (N=34)                 | (N=7)            | (N=1)            |                 |

In the following excerpt the active role of users is realized by the authoritative verbal process in a hypotactic time clause. An authoritative speaker makes use of personalized time summons and refers to Facebook CEO as to the subjected social actor. “Kings” in this piece of writing are opposed to “gods” and symbolize the Facebook leadership, thus, realizing overdetermination through a combination of generalization and abstraction:

11) It’s time we prove that these kings are not gods. It’s time we dethrone Mark Zuckerberg.

The codification of two social roles is manifested in the extract that follows. Facebook executives are activated in relation to the process of “allowing” and beneficialized to the action of “profiting from the sale of ads to users”:

12) Facebook allowed users to purchase ads targeted towards racist interests like “Jew hater” or “Nazi Party”, profiting from the sale of ads to users with those interests.

Table II evidences that “private data seekers”, i.e., people, and organizations that are searching for obtaining personal information, are more passivated (83.33%) than activated (16.67%).

Privacy issues organizations are predominantly (96.77%) presented as active dynamic forces that serve legitimization purposes. The FTC functions as the subject of the verbal process clause and embodies the impersonal authority of regulations, rules and laws protecting American consumers:

13) In 2011, the Federal Trade Commission settled charges with Facebook that the social networking giant “deceived consumers by telling them they could keep their privacy far outnumbers other groups of
analyzed social actors (90.91%), having only several cases of abstract and quality assigned representations. In the following discourse fragments the thorough scrutiny of "privacy organizations" abstraction is given:

14) Any competent regulator must first frame the problem they are trying to address. <…> There are many other potential solutions, but the point of any of them would be to attack the culture of lawlessness in the company and eliminate the motivation to collect data.

<p>| TABLE III. PERSONALIZATION/IMPERSONALIZATION IN THE NEWS REPORTS REGARDING PRIVACY |
|---------------------------------|------------------------------------------|------------------------------------------|------------------------------------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>Personalized</th>
<th>Instrumentalized</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abstrusted</td>
<td>Instrumentalized</td>
</tr>
<tr>
<td>“Users” (N=33)</td>
<td>84.85% (N=28)</td>
<td>15.15%</td>
</tr>
<tr>
<td>“Company management” (N=53)</td>
<td>33.96% (N=18)</td>
<td>66.04%</td>
</tr>
<tr>
<td>“Private data seekers” (N=23)</td>
<td>30.43% (N=7)</td>
<td>39.14%</td>
</tr>
<tr>
<td>“Privacy organizations” (N=33)</td>
<td>90.91% (N=30)</td>
<td>9.09%</td>
</tr>
</tbody>
</table>

The supervisory function of data privacy legislation institutions is identified with privacy agencies through the unspecified quality of “any competent regulator” assigned to them. Public interest litigation actions of privacy organizations are referred to by means of an abstract noun with an attributive adjective “potential solutions”.

Percentage value of personalized users and consumers has a reasonable amount as well (84.85%), leaving the remaining 15.15% to generalized (“public”), categorized as having a permanent role (“visitors”), having “accounts,” associated with individuals (“individually targeted”) abstracted nominations:

15) The bug was discovered in 2018 and patched, but its existence was never disclosed to the public.

16) If a website embeds Facebook features like the “like” button even if the visitor doesn’t have a Facebook account, Facebook is tracking you and shares your demographic data with advertisers.

17) Facebook admitted in 2016 that it was following the direction of the American and Israeli governments and deleting accounts of Palestinians for “incitement” — a standard it did not hold for accounts from other nationalities.

18) Second, the company operates in a market where the incentive is to gather as much personally identifiable data as possible so it can use that data to profit through individually targeted advertising.

The substitution of “individually targeted advertising” for “advertising companies” serves as prototypical case of an abstraction, and what is abstracted is the form of organization of human action aimed at obtaining personal information. Other examples of backgrounding the identity and role of “private data seekers” demonstrate the impersonalized abstracted representations with negative connotations – “spread of fake news,” “distribution of toxic and violent multimedia,” and “an aggressive public relations campaign”:

19) Today, the company is again in hot water for, among other things, misusing private user data, failing to stop the spread of fake news and enabling the distribution of toxic and violent multimedia.

20) In the wake of the Cambridge Analytica scandal, instead of working substantively on internal reforms, Facebook’s policy team went on an aggressive public relations campaign, hiring a right-leaning public affairs group to smear, among others, the left-leaning billionaire George Soros.

In the following contexts agencies searching for private data are presented by means of reference to instruments with that illegitimate actions are carried out:

21) A breach of Facebook’s computer systems left the personal information of 30 million users exposed — searches, locations, names, phone numbers, and more.

22) A White Hat hacker in 2013 found a bug exposing the email addresses and phone numbers of 6 million Facebook users to anyone who had some connection to the person or knew at least one piece of their contact information.

23) One of Facebook’s early advertising programs, Beacon, automatically notified friends when users bought something, without obtaining prior consent.

It can be seen that “private data seekers” are objectivated through instrumentalization, i.e., realized in the text by metonymically reference to “a breach,” “a bug,” an “advertising program” performing the actions of personal data acquisition in that the social actors are engaged.

In the analyzed contexts Facebook responsible executives are little less than twofold abstracted (66.04%) than personalized (33.96%). In most cases of impersonalization the “company management” stands for the Facebook company itself and highlights a form of human action organization governed by impersonal procedures:

24) Facebook admitted it ran experiments on hundreds of thousands of users to test how emotions spread on the site.

As it can also be observed, in the excerpts under study, abstraction contributes to the attachment of connotative meanings to social actors it denotes. Thus, qualities attributed to “royal families”, “the most loathed institutions”, “the ruling caste” are abstracted from their bearers, they serve to interpret and evaluate Facebook “company management” as the most unwanted yet wielding formidable authority personalities:

25) In this gilded age of big tech, Silicon Valley’s royal families have sat in their castles, amassing treasure mined from a mountain of our data, seemingly untouched as they weather scandal after scandal while they offer empty apologies and their power grows.
26) Under Mark Zuckerberg’s failed leadership, Facebook has become one of the most loathed institutions on the planet.

27) But Zuckerberg personally, and the ruling caste at Facebook generally, has shown no appetite for real reform—despite what they say on TV or in front of Congress.

D. Social actions pattern

Social action, according to van Leeuwen, can be interpreted as semiotic if it conveys some meaning as opposed to material action which “has, at least potentially, a material purpose or effect” [25]. Facebook, Google, and Amazon company management featured in the news articles is mostly presented as involved in material transactive actions since they identify and further utilize personal information; it “gave preferential access to user information,” “ran experiments on hundreds of thousands of users,” “used an app called Onavo to learn how long users spent on competing apps,” “was allowing application developers access to personal and behavioral data for millions of Facebook users,” “Google allows companies to target people with adverts,” “Amazon, Google, and Facebook, and up-and-coming troublemakers like Uber were gathering up as much personal data as they could and deploying it in invisible ways,” “was tracking metadata on statuses and posts,” “collected data,” etc. These actions, represented as invasive and having a negative effect on users and public in general, are prevalently instrumental. Minor part of the executives’ actions is manifested as cognitive: they “did discuss selling access to user data to advertisers in 2012,” or as an affective mental process (“wanted to”):

28) Facebook wanted developers to build features like games that would make Facebook more attractive and important to daily life.

In the excerpt that follows conceptual structure of Facebook is opposed to the existing traditions and practices that exist in modern society. The behavioralized action (“is at odds”) of Facebook authorities explicitly differentiates them from other people that are abstracted in the discourse by means of nominations “democracy and basic human rights”:

29) Facebook’s current business model is fundamentally at odds with democracy and basic human rights.

The actions of Facebook managers are repeatedly deactivated and realized by nominalizations and nonfinite forms: “under Mark Zuckerberg’s failed leadership,” “after an admission of wrongdoing, the second step is to restructure the business model of Facebook,” “Mark Zuckerberg faced accusations in a lawsuit,” “Facebook’s reputation on privacy issues is so awful,” “documents also led to Facebook finally agreeing to a hearing in front of the House of Commons,” etc. Objectivation occurs in relation to negatively interpreted and descriptivized actions, having permanent qualities of improper and illegitimate deeds: “people who have little knowledge of or concern for democracy or the dignity of the company’s 2.3 billion users,” “the company has a culture of immorality and lawlessness”. Reactions of Facebook leadership are cognitive and semiotic: they “said,” “confirmed,” “commented,” “disclosed,” “covered up,” “did discuss selling access to user data,” in some cases remedial: “apologized,” “admitted he did wrong.”

The actions of privacy regulation agencies are predominantly activated. They may be characterized as material transactive ones which have a prescriptive function that urges to “attack the culture of lawlessness,” “eliminate the motivation to collect data,” “put aggressive restrictions”, “go to court to enforce,” “move taxpayer money to enforcers,” etc. In the following examples the representation of the regulatory practice downgrades and the priority is given to modality: “the FTC must require something meaningful,” “the FTC could follow the example of the German federal cartel office”. This modality alongside objectivation – “the agency’s fickleness,” “a year-long investigation by the FTC,” “the current FTC isn’t capable of reigning in Facebook” implicate backgrounded criticism of long-term, inefficient and nonpunitive measures taken by the federal agency. The agency and other regulators embody impersonal authority that mandatorily has to impose and uphold legitimacy regarding user data privacy. Hence, the reactions of the FTC are deactivated and negatively descriptivized as not meeting the requirements and expectations of public opinion.

Users, in the main, are engaged in involuntary actions which happen without their direct participation: they “had their passwords stored in plain text,” “had their news feeds manipulated by Facebook researchers,” “had their privacy violated and their trust betrayed,” “had been persuaded”. The processes that users are involved in are deagentialized, they serve as a model for representing naturally occurring and uncontrolled by common people events: “people are eating it up,” “if you played Mafia Wars or Words with Friends through Facebook, you probably revealed personal information about all of your friends,” etc.

The permanent action of a user (“a person visits a website”) is non-transactive and activated, while the “users” are subjected to the description of their personal data “broadcast to tens or hundreds of companies.”

30) Every time a person visits a website and is shown a ‘behavioral’ ad on a website, intimate personal data that describes each visitor, and what they are watching online, is broadcast to tens or hundreds of companies.

If the actions of users and consumers are activated, it occurs in relation to material instrumental actions which are neither goal, nor effect-oriented: “shared their status,” “revealed personal information,” “bought something,” “spent on competing apps,” etc.

IV. CONCLUSION

The present paper analyzes the representation of social roles and their activities involved in the current processes that regard Internet privacy. Van Leeuwen’s socio-semantic inventory allows for thorough investigation of the linguistic factual data in question. All in all, the patterns based on sets of categories of van Leeuwen’s model point out that privacy concern manifests itself in the news reports through the actions and reactions of the following groups of social actors: collectivized users and consumers on the one side, notorious Internet, surveillance, advertising companies on
the other side and intermediary privacy regulators. Internet customers are sympathetically treated and united in their consensus about universal human values and rights. They act as victims who suffer from illegitimate actions of Internet companies which violate privacy laws. In contrast, Internet companies have negative connotations and are portrayed as publicly criticized, illegally deploying personal data entities, social practices of which do not comply with overall expectations. Functionalized privacy organizations are represented as organs imbued with impersonal authority though not executing it in a proper and expected way. More detailed analysis of the problem in relation to the discursive construction of the purposes and legitimation will be bound to shed a new light on the differential representation of social actors involved into current privacy issues.

REFERENCES


[22] M. Kelly, „Facebook could create new privacy positions as part of FTC settlement,” in The Verge, May 2, 2019. URL: https://www.theverge.com/2019/5/2/18526832/facebook-ftc-settlement-privacy-oversight-committee-positions-jobs (date of access: 19.05.19).

[23] S. Hollister, „Read Google’s new Nest privacy promise and tell me if you’re swayed,” in The Verge, May 7, 2019. URL: https://www.theverge.com/2019/5/7/18536184/google-nest-privacy-commitment (date of access: 19.05.19).
