"Alexandru Ioan Cuza" University, Iași
Faculty of Letters
Doctoral School of Philological Studies

Linguistic particularities of the text from Romanian user generated content on online social platforms

- PhD thesis summary -

PhD:

Andrei Sebastian Stipiuc

PhD Supervisor:

Prof.Univ.Dr. Luminița Cărăușu

IAȘI, september 2014

**Linguistic particularities of text from Romanian user generated content
on online social platforms
- Summary -**

I. Context, objectives and research methods

Through the means of the proposed PhD thesis, *Linguistic particularities of the text from Romanian user generated content on online social platforms*, we wish to subscribe to the type of rare synchrony practices in Romanian literature regarding the Romanian online space in order to study the language manifestations of native Romanian speakers who are active on the Web.

Nowadays, the World Wide Web is represented by its most popular applications. The social character of the Web (Berners-Lee 2001: 12), where the users number is increasing year by year, lies in the social networks or in other platforms of the virtual world, where people interact and communicate on a daily basis.

The Web is no longer just a reading place or a place of consumption; it is a place of interactivity and convergence of the daily human activities (Morris 1996). Given this fundamental social context, language still plays its essential role in user interaction in local, national or even global communities.

Our paper is traditionally based on the computer mediated communication setting, although we are trying to approach, through the perspective of the digital humanism that we need to discuss, what David Crystal was saying by using the term *Internet Linguistics* (2004) about the study of Internet language.

The PhD thesis is oriented towards the Web innovations of the latter years – the UGC (User Generated Content) platforms and their Romanian users, mainly towards the members of social[1] networks (Gumperz 1996: 27-38). Although there is an entire terminology debate, we preferred the term "social network", which we have adapted out of reasons concerning the sociolinguistic approach of our paper, keeping the term as it was created by Gumperz.

We followed the way in which the variations of the written language are manifesting in the online world in its permanent adjustment to the characteristics of these platforms. We took into consideration the features of the linguistic version of written language (text): graphic features (aspects regarding typography, design, layout, illustration, colour) and orthographic ones (the written system of the Romanian language on these platforms: alphabet, capital letters, diacritics, punctuation, emphasis – cursive, bold, underlined), the lexical-semantic features (vocabulary items, terms that have appeared due to the penetration of online activities into daily life, idioms) and the morpho-syntactic ones, the speech characteristics through which personal narration is built on Facebook in particular, methods in which online politeness is manifesting or its

---

[1] The social perspective that we have chosen for our study regards Gumperz's definitions that identify and qualify the existent relationships between the members of the linguistic communities.

**Linguistic particularities of text from Romanian user generated content
on online social platforms
- Summary -**

infringement possibilities, marking of positive and negative feelings, code switching depending on cultural isolation, immigration of users to other countries and the adoption of the local language, the continuous expansion of "local English languages" (Schneider 2011: 336), the practices through which computer slang manages to influence the norm of the offline world. There is a series of common linguistic characteristics for both synchronic and asynchronous communication (usage of the emoticon, abbreviations, shortened writing but also an inclination towards verbal informality), which brings written language even closer to the features of spoken language.

We focused the research around status updates on Facebook, because there is a consensus that these updates, no matter their origin or subject, offer indications upon the state of mind of each individual that shares them. The reasons of distributing (the *share* action) are of utilitarian, informational, militant and mostly autobiographical nature, and users sending out information understand it better when they re-transmit it. It has been found that 80% out of published messages on Facebook have biographical content. Linguistically, the intertextuality of the environment and the resemblance with publishing regarding phrases and idioms as part of repetitive speech are clear. In speech, there are questions, as a feature of interaction,

**Linguistic particularities of text from Romanian user generated content
on online social platforms
- Summary -**

pre-sequences with which the availability of the receiver is checked, memorial writing.

We have also inferred an evolution of communication through linguistic-imagistic means, by the use of *memes* (combining text and image stupidly on purpose, with a high virality degree), but also through the lines provided with hypertext that send to images – as an answer marking a personal opinion following a selection from a huge data base of potential answers.

The linguistic features of written content generated by Romanian users, that we have identified among social networks and of which we compiled the corpus of this paper, have been grouped in the following manner:

1. Orthographical features:

   Abbreviated writing; Acronyms writing; Alphanumeric writing, by substituting letters with graphically similar numbers; Unconventional writing, using "@" for email addresses with the purpose of marking its appurtenance to the online space;

   Telegraphic writing (short sentences, rapidly typed, with omissions to punctuation or orthography, writing without spaces after punctuation marks, lowercase writing after punctuation marks or in the beginning of the sentence, writing without diacritics);

Writing or using symbols and emoticons to reproduce the paralinguistic effects in writing (emotions, expressiveness), which otherwise are found only in face to face communication;

Writing by repeating consonants or vowels in order to emphasize a feeling.

2. Some lexical features of technical slang or which define interface elements of online platforms that have left the online world and entered daily use.

3. Interface syntax, that includes the way in which language is used to provide written means of communication to the user (for writing own texts, sharing different content online – hyperlink, images). Using the "#" sign in order to create tags (*hashtags*) that would subscribe the intervention to a semantic area around the same subject. Using Facebook tools in order to complete predefined fields through which a user can mark certain moments in their personal life.

4. Code switching, in which Romanian language interferes with a foreign language (usually English, but we can also encounter French and Italian) or in which a language different from Romanian is selected to mark the membership to the new socio-cultural community, usually in the case of

**Linguistic particularities of text from Romanian user generated content
on online social platforms
- Summary -**

Romanian immigrants. For most Romanian users, English became *lingua franca* in the online medium.

5. Text + image language (*LOLcats, meme*), turned from its concept and basic purpose and now used for plastic and playful-parody expression of daily anxieties or to ironically answer some comments.

6. Compliance or violation of online politeness, whose rules are consisted into a *netiquette* that some users follow, while others do not.

7. Linguistic games, within which invented lexicon stands out, the parody of grammar mistakes, phraseological units as part of repeated speech, through which users are manifesting their inventiveness and linguistic creativity.

8. The metalinguistic features, in which language is discussed in general, and in which a stand is taken towards wrong use of some terms and against those who are not familiar with the grammatical norm.

9. The pragma-linguistic features of personal narration, the defined everyday ego through memorial interventions or through texts celebrating daily or annual social rituals: Romanian and international holidays, personal celebrations etc.

Although we have proposed a descriptive study in our analysis, we also took into consideration the materials that would not necessarily contain negligent or incorrect articulation by grammatical standards, those materials that do not reveal misuse of language and we subjected to analysis only the relevant materials according to the point of view of alterations imposed by the environment. The examples for the corpus of the doctoral practice were gathered over two years (2012-2014). Following a rigorous selection, we published the examples that illustrated, with no doubt, the linguistic features of written Romanian user generated content online.

Besides public materials, the hereby corpus is mainly constituted of texts gathered from two different Facebook accounts, from which a number of approximately 1000 unique users was weekly observed. We opted for the censorship of the surname or of the second user name in order to offer privacy to the authors of the published texts. The chosen examples are written by Romanian users, living in Romania (from Iasi, Bucharest, other cities in the Moldavian region), but also from the Republic of Moldova.

A small, yet important number of users out of all the observed ones is now established over boarders (The United Kingdom, Italy, France), and we have specified this aspect for every subject that needed this mention. Approximately 45% out of the followed users are students (20-25 years old) with higher education

(Bachelor's degree, Master's degree, PhD), 30% are young people with ages between 25 and 30 years old and the rest of 25% are older than 30.

II. Research structure

In the first part of the paper (Chapters I-III), we have established the general research frame of the language as a fundamental concern in analysing computer mediated communication, mentioning the pressure coming from the environment that impacts natural language. We have noted the main foreign studies, as well as the Romanian contributions to this field. The new attempt was looking for and finding a common root (regarding text messages - SMS) with the linguistic variety online, the subject of our research. This was also the moment of the first endeavours to fit our study into a subdomain initially promoted by David Crystal, and others[2], called *Internet Linguistics*.

The second part of the paper (Chapters IV-V) establishes the methods used to study and build the corpus of the linguistic variety specific to the Internet. We turned to the concepts and methodology of sociolinguistics because the online world and namely the Web

---

[2] The effervescent activity in this domain happened around the publication of the *Journal of Computer-Mediated Communication* (JCMC), http://jcmc.indiana.edu/.

**Linguistic particularities of text from Romanian user generated content
on online social platforms
- Summary -**

have an essential role, its users – Romanian speakers, being connected with the study themes of sociolinguistics such as linguistic variety, users/speakers community, code switching, the relationship between Romanian and English and other international languages inside multilingualism, politeness in computer mediated communication.

The proper analysis of written interventions (asynchronous) of Romanian users has been developed in the third part of our paper (Chapter VI), and the examples were systematized into 9 main categories (in Annex 1).

The linguistic study of the main social network (Facebook) and some incursions in the leading micro-blogging platform (Twitter) has identified sociolinguistic and pragmalinguistic themes, like deviations from politeness norms (netiquette), special online orthography, code switching, linguistic games, communication habits, metalinguistic messages, language in text and image, interface syntax and personal narration.

III. Conclusions

For today's philologist, the online corpus can be an extremely fertile research field, as some digital tools and methods that they can choose from, starting with the *mainstream* ones (usually used for office work, organizing personal data or having an

educational purpose) which can be turned to research and finishing with specialized applications, created by other researchers and made available to the scientists community are numerous. Online user generated text offers new analysis possibilities and can be approached from different[3] directions, in terms of linguistics, communication theories or pragmatics.

Supporting the above mentioned idea, further research should consider the computerization of the current format of the selected corpus (its transformation and structuring into processing data), in order to have it published and open to more research options either for the theoretical linguists or computer linguists.

IV. Notes:

Berners-Lee 2001: Tim Berners-Lee, Max Fischetti, *Weaving the Web*, Barnes & Noble, p.12.
Burdick 2012: Anne Burdick et al., *Digital Humanities*, Massachusetts Institute of Technology.
Crystal 2003: David Crystal, *The scope of Internet linguistics*, http://www.davidcrystal.com/DC_articles/Internet2.pdf .
Crystal 2004: David Crystal, *Language and the Internet*, Cambridge University Press.

---

[3] Supporting this idea, we have proposed the article called *"Tools and digital methods in computer-mediated language research. The relevance of the online corpus"*, under review ((*1st International Conference Perspectives in the Humanities and Social Sciences: Hinting at Interdisciplinarity,* Iaşi, Romania, May 23rd-24th 2014).

Gumperz 1966: John Gumperz, *On the Ethnology of Linguistic Change*, în: W. Bright (Ed.), *Sociolinguistics*, The Hague, Paris, p. 27-38.

Morris 1996: Merrill Morris, Christine Ogan, *The Internet as a Mass Medium*, Indiana University, Journal of Communication.

Schneider 2011: Edgar W. Schneider, *Colonization, globalization, and the sociolinguistics of World Englishes* în Rajend Mesthrie (Ed.), *The Cambridge Handbook of Sociolinguistics*, Cambridge University Press,, p.335.

V. Selective bibliography

Androutsopoulos, Jannis; *Sociolinguistics and computer-mediated communication*, în Journal of Sociolinguistics 10/4, 2006.

Baron, N. S.; *Letters by phone or speech by other means: The linguistics of email, Language and Communication*, 1998.

Berners-Lee, Tim; Fischetti, Max, *Weaving the Web*, Barnes & Noble, 2001.

Biber, D., *Spoken and written textual dimensions in English: Resolving the contradictory findings*, *Language* 62/2, 1986.

Bidu-Vrânceanu, Angela et. al., *Dicţionar de ştiinţe ale limbii*, Editura Nemira & Co., Bucureşti, 2005.

Brown, Penelope; Levinson, Stephen C. *Politeness. Some universals in language use*, Cambridge University Press, 2004.

Bruthiaux, P., *Predicting challenges to English as a global language in the 21st century. Language Problems and Language Planning*, 26(2), 129–157, 2002.

Bucholtz, Mary; Hall, Kira; *Twenty years after Language and Woman's Place*," în Hall and Bucholtz (eds.), 1995.

Cameron, Deborah, *Language, gender, and sexuality: current issues and new directions*, *Applied Linguistics* 26, 2005.

Chomsky, Noam, *Cunoaşterea limbii*, trad. Alexandra Cornilescu, Ileana Baciu and Taina Duţescu Colibar, Editura Ştiinţifică, Bucureşti, 1996.

Cmeciu, Camelia Mihaela, *Strategii persuasive în discursul politic*, Universitas XXI, Iaşi, 2005.

Coșeriu, Eugeniu, *Omul și limbajul său*, coord. Dorel Fînaru, coord. trad. Dumitru Irimia, trad. Eugenia Bojoga, Florin Bratu, A. Covaciu, Editura Universității „Alexandru Ioan Cuza Iași", 2009.

Crystal David, *Language and the Internet*, Cambridge University Press, 2004.

Cvasnîi Cătănescu, Maria, *Retorică publicistică. De la paratext la text*, Editura Universității din București, 2006.

Danet, Brenda; Herring, Susan C.; (Eds.), *The Multilingual Internet. Language, Culture, and Communication Online*, Oxford University Press, 2007.

Dimitrescu, Florica, "Dicționar de cuvinte recente", second edition, Editura Lagos, 1997.

Dimmick, J.W.; Kline, S.; Stafford, L.; *The gratification niches of personal email and the telephone: Competition, displacement and complementarity*, Communication Research, 2000.

Duranti A. and Goodwin, C. (Eds.), *Rethinking context: Language as an interactive phenomenon*. Cambridge: Cambridge University Press, 1992.

Edwards, John, *Multilingualism,* London: Routledge, 1994.

Fishman, J. A.; *The new linguistic order*, *Foreign Policy*, Winter, 26–40, 1998.

Franklin, Ursula, *The Real World of Technology*, Anansi, 1999.

Geertz, C; *Interpretation of cultures*. New York: Basic Books, 1973.

Genette, Gérard; Marie Maclean, *Introduction to the Paratex*t în New Literacy History, Vol. 22, Nr. 2, *Probings: Art, Criticism, Genre* (Spring, 1991).

Gibbs, D.; Krause, K-L (Eds.), *Cyberlines 2.0: Languages and Culture of the Internet*, Melbourne: James Nicholas Publishers, 2007.

Goffman, Erving, *Interaction Ritual. Essays in Face-to-Face Behaviour*, Garden City, New York, Doubleday et. Co, 1967.

Grinter, D.E.; Eldridge, M.A.; *y do tngrs luv 2 txt msg?*, 2001, in W, Prinz, M, Jarke, Y, Rogers, K, Schmidt & V, Wulf (eds), *Proceedings of the seventh European Conference on Computer-Supported Cooperative Work*, Netherlands: Kluwer Academic Publishers.

Gumperz, John, *Types of Linguistic Communities*, p.116, in: Liliana Ionescu-Ruxăndoiu; Dumitru Chițoran, *Sociolinguistics – Current orientations*, Teaching and Pedagogic Publishing House, Bucharest, 1975.

Haugen, E., *The ecology of language* in Dil, A. (Ed.), *The Ecology of Language: Essays by Einar Haugen*. Stanford University Press, 1972.

Herring, S.C. (Ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. Amsterdam: Benjamins, 1996.

Hoarță-Cărăușu, Luminița, *Teorii şi practici ale comunicării*, Editura Cermi, Iaşi, 2008.

Hymes, Dell; *Foundations in Sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press, 1974.

Ionescu-Ruxăndoiu, Liliana; Chiţoran, Dumitru; *Sociolingvistica - Orientări actuale*, Editura Didactică şi Pedagogică, Bucureşti, 1975.

Johnstone, Barbara; *The Linguistic Individual. Sel-Expression in Language and Linguistics*, Oxford University Press, New York, 1996.

Kachru, Braj B. (Ed.), *The Other Tongue: English across Cultures*. Urbana and Chicago: University of Illinois Press. 2nd ed. 1992.

Kleinman, Zoe; *How the internet is changing language*, 16 august 2010, BBC NEWS TECHNOLOGY, http://www.bbc.co.uk/news/technology-10971949 .

Kramer, Adam D. I.; Chung, K.; *Dimensions of Self-Expression in Facebook Status Updates*, Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media Barcelona, Catalonia, Spain, July 17 – 21, 2011.

Labov, William; *Sociolinguistic patterns*. Philadelphia: U of Pennsylvania Press, 1972.

Levinson, Paul; *Digital McLuhan: A Guide to the Information Millenium*, Routledge, London, 2004.

Lewis, M. P., *Ethnologue: Languages of the World*. 15th edn. Dallas, TX: SIL International, 2009.

Mazzarella, Sharon M. (Ed.), *Girl Wide Web: Girls, the Internet and the Negotiation of identity*. New York: Peter Lang.

McLuhan, Marshall; Powers, B. R.; *The Global Village*, NewYork: Oxford University Press, 1989.

McQuail, Denis; Windahl, Sven; *Modele ale comunicării pentru studiul comunicării de masă*, Comunicare. ro, Bucureşti, 2004.

Media Fact Book 2013, Initiative Media S.A., Bucharest, 2013.

Mesthrie, Rajend (Ed.), *The Cambridge Handbook of Sociolinguistics*, Cambridge University Press, 2011.

Morris, Merrill; Ogan, Christine, *The Internet as a Mass Medium*, Indiana University, Journal of Communication, 1996.

Muñoz-García, Óscar; Navarro, Carlos; *Comparing user generated content published in different social media sources*, lrec-conf.com, mai 2012.

Olson, D. R., Torrance, N. şi Hildyard, A. (ed.), *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing*. Cambridge University Press, 1985.

Parks, M.R.; Floyd K., *Making friends in cyberspace. Journal of Computer-Mediated Communication*, 1 (4), 1996.

Peters, Nikki; *Social Media Language, Taking Over the Oxford English Dictionary*, 21 august 2011, socialmediatoday.com, http://socialmediatoday.com/marketmesuite-app/340584/social-media-language-taking-over-oxford-english-dictionary .

Sapir, Edward; *Language*. New York: Harcourt, Brace, and World, 1921.

Saussure, Ferdinand de, *Curs de lingvistică generală*, Editura Polirom, Iaşi, 2005.

Savignon, Sandra J.; Roithmeier, Waltraud; *Computer-mediated Communication: Texts and Strategies* în CALICO Journal, 21 (2).

Schneider, Edgar W., *Colonization, globalization, and the sociolinguistics of World Englishes* în Rajend Mesthrie (Ed.), *The Cambridge Handbook of Sociolinguistics*, Cambridge University Press, 2011.

Soukup, C., *Building a theory of multimedia CMC: An analysis, critique and integration of computer-mediated communication theory and research*. New Media & Society 2(4), 2000.

Sterne, Jonathan; *Thinking the internet: cultural studies v. the millenium*. În S. Jones (ed.), *Doing Internet Research: Critical issues and methods for examining the net* (pp. 257-288), London, Sage, 1999.

Steuer, J. (1992). *Defining virtual reality: Dimensions determining telepresence*. Journal of Communication 42(4), 73-93.

Tannen, Deborah (ed.), *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: Ablex.

Thurlow, Crispin; Lengel, Laura; Tomic Alice, *Computer Mediated Communication. Social Interaction and the Internet*, Sage Publications Ltd., London, 2004.

Walther, J.B.; *Interpersonal effects in computer mediated interaction: a relational perspective*. Communication Research, 19, 52–90, 1992.

Wittkower, D.E. (Ed.), *Facebook and Philosophy*, Open Court, Chicago and La Salle, Illinois, 2010.

Wong, Anette, *Cyberself: Identity, Language and Stylisation on the Internet*. in D. Gibbs & K. Krause (Eds.), *Cyberlines 2.0: Languages and Cultures of the Internet*. Melbourne: James Nicholas Publishers, 2007.

Iaşi, September 2014