Language change online

Linguistic innovations in Russian induced by computer-mediated communication

Aleksandrs Berdicevskis

Dissertation for the degree philosophiae doctor (PhD)

at the University of Bergen

2013

Dissertation date: 29.01.2013

Scientific environment

This work was carried out while I was a PhD student at the University of Bergen, Department of Foreign Languages (August 2009–February 2013, funded by the University of Bergen) and a member of the "Future of Russian: Language Culture in the Era of New Technology" project (August 2009–July 2012), led by Professor Ingunn Lunde and funded by the Research Council of Norway.

Acknowledgments

While working on the dissertation I received help from many people. First of all, I would like to express my gratitude to my supervisor, Ingunn Lunde, who conceived the "Future of Russian" project, secured generous funding and guided the project enthusiastically and wisely. Ingunn supported all my initiatives, including those that were not directly relevant to the dissertation.

The two postdoctoral members of the core group of the "Future" project, Martin Paulsen and Ellen Rutten, were also always ready to help, to explain, to teach, to comment. I would also like to thank other "futurants" for being a great international team. I am particularly grateful to Lara Ryazanova-Clarke for my three-month stay in Edinburgh in 2011 and to Vera Zvereva for her frequent and inspiring visits to Bergen.

The dissertation benefited from discussions with guest lecturers invited by the "Future" project, in particular Jannis Androutsopoulos, Vladimir Plungân and Nina Mečkovskaâ. Likewise, I am grateful to guest lecturers and/or commentators invited by the PhD Research School in Linguistics and Philology, in particular Jane Stuart-Smith, Chris Butler and Paul Kerswill. Thanks must also go to school leaders Victoria Rosén, Kari Haugland and Øivin Andersen and school coordinator Martin Paulsen.

Discussions with colleagues both within and outside UiB were of great importance to me. I would like to thank Lars-Eric Christer Johansson, Aleksandr Piperski, Edit Bugge, Bodil Moss, Boris Iomdin, Arkadij Avdohin, Nataliâ Aralova, Vladimir Belikov and Evgenij Šaul′skij for the various forms of help they gave me. Special thanks go to Viktor Sonkin, who translated Chapter 3 into English, and to Il′a German, who developed software which allowed me to store and manage some of my data in a very convenient way.

I spent several months teaching, and I am grateful to the colleagues at the Russian department who made it possible and with whom I worked: Margje Post, Brita Bryn, David Pineda, Malvin Fonnes and Lillian Helle. Thanks to Kåre Johan Mjør and Elisabeth Tiselius for being great office mates.

I am grateful to the Research Council of Norway and the University of Bergen for funding, and to the latter institution also for providing a great environment for a PhD student. The important figures in this environment's academic constituent have already been mentioned above, but there are also those who ensure the smooth work of administrative mechanisms, and thanks goes to all of them, in particular to Arve Kjell Uthaug, Kristin Skogedal and Anne Hestnes.

Living in Bergen, I was very lucky to have amazing landlords, first Jardar Østbø, and then Marie and Erik Valevatn.

Kyrill Dissanayake quickly, thoroughly and efficiently erased the traces of Russian from my English. If any instances of imperfect English remain anywhere, these are the results of last-minute changes that evaded Kyrill.

Finally, the work on the dissertation would not have had much sense were it not for my wife Tanja, daughter Julia and son Jan, and would have never become possible without the support, love and foresight of my parents, Sergej Berdičevskij and Elena Averjanihina.

Foreword

While this dissertation uses data from Russian language, every effort has been made to ensure it is accessible to a reader who does not have any knowledge of Russian. All examples are provided in Cyrillic and in Latin transliteration. Apart from translation, glossed or explicit commentaries are provided where necessary.

I use the ISO 9: 1995 Romanization system (equivalent to Russian state standard GOST 7.79—2000, system A). While this system is seldom used in Western publications, it has several important advantages over the most popular system, that of the Library of Congress. First, the LC system is not reversible, since it does not ensure biunique correspondences between Cyrillic and Latin characters. This applies both to the "simple" diacritic-free version (which does not even distinguish between *и* and *ĭ*; *я*, *йа* and *иа*; *ц* and *тс*; etc.) and the "complex" one, which makes extensive use of diacritics and is supposed to be unambiguous. Nevertheless, it is not: *shch* can correspond both to *щ* and to *шч*. My examples abound with deviations from the norm, and in some cases orthographic deviations are the focus of interest (in Chapter 6, for instance), and so it is important to avoid any loss of information in the transliteration of examples. For the same reason, it is convenient to use a system in which one Cyrillic character is always represented by one Latin character, as is the case in ISO 9, not a combination, as is the case in LC. Moreover, ISO 9's diacritics are more convenient to use than LC's two-character ties. Finally, ISO 9 is more standardized: for each Latin character, its Unicode number is defined (nothing similar occurs in LC). Thus, for a linguist interested in the exact and simple representation of Cyrillic examples, ISO 9 is a perfect option, while LC is one of the least suitable. For consistency, ISO 9 is also used for bibliographical references.

Glosses follow the Leipzig Glossing Rules with some modifications, explained where necessary.

Abstract

In this dissertation, I study the role of computer-mediated communication (CMC) in language change. More precisely, I address the question of how speakers adapt their linguistic behavior to the constraints and affordances of the new communication channel, and whether changes in their behavior lead to the emergence of any innovations in the Russian language.

In Chapter 2, I compare the linguistic variation in two communication channels: e-mail and chat, demonstrating that for some variables there exist significant differences. I show that these differences can be explained by differences between the constraints and affordances of the two channels. In Chapter 3, I analyze a recent morphosyntactic innovation in Russian, the so-called "orthographic neuter". I provide evidence for the fact that the successful diffusion of this innovation occurred in CMC, and would have been impossible without CMC. Chapter 4 has two aims. First, I describe some aspects of the so-called written turn, the trend towards greater autonomy and salience for written speech, reinvigorated by CMC. Second, I use this description as an example when addressing the well-known *actuation problem* of sociolinguistics and analyzing the role of CMC in actuating language change. I claim that most of the influence that CMC exerts on language change occurs not at the stage of individual speaker innovations, but at the stage of diffusion. In other words, CMC changes the selective forces that operate in language and determine which speaker innovations are likely to diffuse across society and which are not. In Chapter 5, I study how users play with conversation structure in quasi-synchronous CMC (chats and instant messengers). I use the methods of conversation analysis to elicit novel patterns of playful conversational behavior. Further, I analyze these playful, "anomalous" patterns in order to draw some conclusions about the "normal" management of conversation in quasi-synchronous CMC. In Chapter 6, I provide a quantitative description of the diachronic development of an online slang. The slang in question is the so-called "Olbanian language", an orthographic anti-language,

probably the most widely-known phenomenon in Russian CMC. I show how the frequency of use of anti-normative forms changes over a period of eleven years, both within and outside the Olbanian community.

In the introductory Chapter 1, I discuss methodological issues, in particular, the danger of overestimating the role of CMC and ignoring other factors (i.e. the pitfalls of technological determinism). Further, I describe the general model of language change I adopt — "language as a complex adaptive system" — and provide an overview of the results of all five chapters from the standpoint of this model.

List of publications

Chapter 2 is a revised version of:

Berdicevskis, Aleksandrs. 2011. “E-mail vs. chat: The influence of the communication channel on the language.” In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference “Dialogue”* 10 (17), 84–93. Moscow: RGGU.

Chapter 3 is a revised and translated version of:

Berdičevskij, Aleksandr. 2012. “«Orfografičeskij» srednij rod: grammatičeskaâ innovaciâ v âzyke russkogo Interneta.” [The “orthographic” neuter: a grammatical innovation in Russian internet language] In *Variativnost' v âzyke i kommunikacii* [Variation in language and communication], edited by Lûdmila Fëdorova, 51–72. Moscow: RGGU.

Chapter 4 is a slightly revised version of:

Berdicevskis, Aleksandrs. “The written turn and its role in actuating language change.” In *Digital Russia: The Language, Culture, and Politics of New Media Communication*, edited by Michael Gorham, Ingunn Lunde and Martin Paulsen. Forthcoming.

Chapter 6 is a slightly revised version of:

Berdicevskis, Aleksandrs. “Rise and fall of the Olbanian language: diachronic description of an online slang.” In *Digital Russia: The Language, Culture, and Politics of New Media Communication*, edited by Michael Gorham, Ingunn Lunde and Martin Paulsen. Forthcoming. (This text will most probably be a section in a joint chapter with Vera Zvereva. The text itself, however, is single-authored).

Contents

I

1. Introduction

1. Three common assumptions about language and the Internet

A renowned typologist once asked me what the subject of my dissertation was. I answered that it was "Russian online", and said yes when she followed up by asking: "And do you take your data from the Internet?". The surprised reaction of my interlocutor was: "But you can find *anything* there!".

This statement reveals at least three implicit assumptions that many linguists and non-linguists share about the Internet and language. First, the Internet is viewed as something that is separate and autonomous from "real life". Second, this independent territory is assumed to have its own language, also separate and different from language offline, and in this language "anything" can happen. Third, this "anything" is considered to be of little interest to researchers of language.

Before presenting the content of the dissertation, I would like to challenge these assumptions. First, the Internet, and computer-mediated communication in general (CMC), are interwoven into the daily lives of many people to such an extent that it is difficult to separate what is online and what is offline. The same applies to linguistic practices. People read classic literature and write scholarly articles online, and discuss their recent battle in an MMORPG (massively multiplayer online role-playing game) or troubles with installing a new IM (instant messaging) client offline. True, the channel where the communication occurs does influence language, and that influence is the focus of this dissertation. It is, however, important to realize that the channel is just *one of the factors* that shape the linguistic properties of communication, that CMC is not a single homogeneous channel, and that practices typical of one channel can be borrowed by another one (i.e. from CMC to oral speech, as will be shown in Chapter 4).

In the dissertation, I will be contraposing online and offline language, but it is important to remember that this is a *simplification*, carried out consciously in order to

address a specific question: how does our linguistic behavior change when we use the now familiar but still recent communication channel — CMC?

Second, while the "language of CMC" (in quotes, since there is no single homogeneous CMC language) does indeed abound with deviations from the norms of the standard language, it is not true that *anything* can be found there. One can find innovations which are indeed surprising, but it is wrong to state that the linguistic behavior of its speakers is completely chaotic and devoid of any patterns. Readers familiar with the history of the Russian language can see some parallels here with the study of the Novgorod birchbark letters (Zaliznâk 2004). For several decades, researchers were unable to read most of the excavated letters correctly. Faced with an incomprehensible fragment (and such fragments were numerous), they presented it as an error committed by an illiterate scribe or by a foreigner, and assumed that *any errors* were possible. This assumption allowed them to interpret the sources of ambiguity in whichever way they wished, and this approach yielded very little knowledge about the language of the birchbark letters and their contents. When linguists who abandoned this approach (Zaliznâk and his colleagues) began their study of the birchbark letters, they quickly discovered that the texts were indeed rich in deviations from previously known varieties of Old Russian, but most of these deviations fitted into a perfect system, the so-called Old Novgorod dialect. Reading the letters while keeping the rules of this system in mind produced surprising results. Many of the letters, which had previously been unreadable, were given clear and unambiguous interpretations, and these interpretations were consistent with each other. The Old Novgorod dialect was given a detailed description. Ultimately, it turned out that the letters contained only a few random errors.

The situation in the Russian Internet is, of course, different (the average Internet user is less literate than the average Novgorod scribe, and random errors are more frequent), but what is common to both is the need to abandon the "anything is possible" approach. As soon as we do so, we face several important questions. If something is possible and something is not, what is the reason for that? Why did the

innovations that can be observed arise? Which trends in language change can be identified? Which specific cases are of most interest? Can we attempt to predict which innovations are likely to emerge in the future and which are not?

Third, the variation exhibited by language in CMC is of scholarly importance. Consider the questions outlined in the previous paragraph. If we manage to answer them, we can better understand the role of the communication channel in language change, and, thus, language change in general. Moreover, innovations born out of CMC are interesting per se, given their frequency and salience for the general public.

There is no reason to dismiss the observed innovations simply because a researcher finds them too strange, or "illiterate", or simply unpleasant. The use of the word "unpleasant" is not a joke: consider the following quote about the *âzyk padonkaff* ('language of scumbags'), a popular anti-orthographic slang in Russian CMC (discussed in detail in Chapter 6): "It is really unpleasant to spend time on 'scumbags' that could have been used to study texts of much higher quality and cultural significance, but, as the saying goes, one must know one's enemy." (Sidorova 2006, 34, translation mine). For contemporary Russian culture, at least for part of it, "scumbags" are of great significance, whatever assessment one makes of the quality of texts produced by them. The categories of "enemies" and "friends" are hardly appropriate to linguistic research.[1]

The main reason why Sidorova considers scumbags enemies is their allegedly negative influence on the literacy of Internet users. The affect of CMC on literacy is a legitimate and important research question. While Sidorova provides no evidence to support her claim, some empirical research in this direction is being done, cf., for instance, Plester and Wood 2009. I do not address this question and do not classify the innovations as "good" or "bad", "harmful" or "useful".

[1] This judgmental approach seems to be more typical for Russian tradition than for Western traditions.

To sum up, CMC is now an inherent part of our life. Presumably, when we use it, we adapt our linguistic behavior to the properties of the communication channel, and this influences the linguistic properties of our speech in a non-random way. The nature and mechanisms of this influence, their cognitive, social and structural aspects, are of great interest to linguists.

In this dissertation, I intend to take a step towards a better understanding of the role of CMC in linguistic change. The potential field of research is extremely broad. In the following five articles, I investigate various aspects of that field. The basic research question I address in all the articles is: "how does CMC influence language change", or, more narrowly, "are there any linguistic innovations in Russian that could not have emerged without CMC?".

A fair number of studies have investigated similar questions, focusing on CMC as a factor in language variation and change. There is, however, no unity among researchers about what the answers to these questions are, or about what methods should be used to look for them, or even whether the questions themselves are worthy of attention or, in actual fact, legitimate. I will use the next section of this introductory chapter to explain why these questions are important and valid, as well as to address some general methodological issues. In the third section, I continue the discussion of methodological frameworks. In the fourth section, I provide an overview of the results of all five articles.

2. Legitimacy of research questions and methods

2.1. The "anti-deterministic" criticism

Within the field of language and CMC studies, approaches such as those suggested in the previous section have been criticized, most explicitly by Androutsopoulos (2006; 2008; 2011a), and also by Georgakopolou (2006), Squires (2010, 461), Dahlberg (2004) and others. In particular, formulating questions in the way I did above means running the risk of exposing one's work to labels such as *technological determinism*,

obsolete *first-wave linguistic CMC studies* (Androutsopoulos 2006) and a *reductionist approach* (Dahlberg 2004).

In this section, I discuss some methodological issues in my work, assessing whether this criticism applies to the solutions I produce. I argue that my research questions are important, interesting and can be addressed with valid scientific methods that produce falsifiable results.

I start by defining some key notions in section 2.2. In section 2.3, I explain my theoretical position, as well as using either theoretical reasoning or empirical findings to address the most important points of criticism that might potentially apply. In section 2.4, I outline how more specific criticism of the "channel influence" approach can be productively taken into account by improving the methodological framework. In Section 2.5, I draw conclusions.

2.2. Key notions

Communication channel

By communication channel I understand the transmission medium where the communication occurs. The word *medium* is probably used more often in this meaning, but is more ambiguous (and forces one to wonder which form to choose each time one has to use it in plural), so I opt for a more narrow *channel* (originally borrowed from information-theoretic models of communication).

The degree of precision with which a channel is defined can vary. For instance, we can speak of face-to-face (FtF) communication, or we can choose to go into greater detail and specify whether this is vocal or hand-signed communication, or whether this is vocal communication occurring under "normal" circumstances or in complete darkness near a roaring waterfall. The roaring waterfall is mentioned in all seriousness, cf. the hypothesis that noise from heavy machinery may have been one of the factors that limited linguistic input by immigrant laborers in northwestern

European countries in the 1960s and thus diminished their chances of reaching proficiency in a second language (Matras 2009, 73).

Similarly, making a strong generalization, we can speak of CMC as a communication channel, but it is obvious that many different subchannels can be singled out. An online newspaper is different to a blog, which is different to e-mail, which is different to chat, which can also be different, depending on which type of chat we are speaking about.

My focus is mostly on *linguistic innovations*, not general descriptions of channels. In other words, I aim not to describe anything like the "language of chatrooms", but to investigate the origin and diffusion of some innovations that have emerged in any of the channels that fall under the 'CMC' label. Most of the data that I use come from the set of CMC registers[2] which can be described as *visual*, *text-based*, *public* and *unedited*, i.e. chats and instant messengers, forums, blogs, imageboards etc.

It can be noted that, while the characteristics *visual*, *text-based* and *public* can be viewed as physical properties of the channel, *unedited* refers instead to its social use. It is technically possible, albeit improbable, that, in a multi-user chat, speakers would carefully plan each utterance and then have it polished by a professional editor, although in reality this can hardly happen. Being provided with an easily accessible written channel, people are eager to use it for high-speed informal communication, and this combination of physical features and their social appropriation results in huge amounts of spontaneous written speech, something which has never existed before on such a scale.

In some cases, it can be difficult to disentangle physical properties from social ones, and this is not something I intend to do. What I am interested in is how a new channel

[2] Register is the variety of language used in a certain situation. The situation is defined by a number of parameters (addresser, addressee, place, time, subject matter, communication channel etc.) (Biber 1994).

affects the process of language change by virtue of a certain set of properties and the new opportunities it offers to speakers.

Constraints and affordances

At the level of an individual user, it is often convenient to speak of the properties of the channel in terms of *constraints* and *affordances*[3]. The concept of *affordance* was coined by Gibson (1977), for whom an affordance was something that the environment offers an animal or human, a specific opportunity for action. While an affordance can be viewed as something that *can* or *cannot* be done, the term *constraint* is used more often now to describe the latter.

For instance, a typical instant messenger (IM) provides its user with the affordance of receiving a message from the interlocutor as soon as it is typed and sent (cf. snail mail, where message delivery takes several days), but imposes the constraint of not being able to see the message before it is sent (cf. FtF conversation, where anything uttered can immediately be heard). Of course, we might also say that a user has the affordance of not showing their interlocutor the message until it is finished, something we do not get in oral speech. Thus, the choice of label — affordance or constraint — can be rather arbitrary, and depends on one's viewpoint. In most cases, however, it is quite clear whether what is being discussed is something that *can* be done (but is impossible in other channels) or something that *cannot* be done (but is possible in other channels).

The Gibsonian view of an affordance as something that is absolute, objective and not dependent on the observer was subsequently criticized (see overview in Hutchby 2001, 26-30 et passim, Lee 2007, 225-227). Norman (1990, 1999) offered the

[3] Instead of *affordances*, some authors choose to speak of *enablements* (Hård af Segerstad 2002, 10-11), implying the same or nearly the same thing.

concept of *perceived affordances*, mental interpretations of what an environment offers a user, based on the user's knowledge and experience.

For instance, mean utterance length in IM is usually rather low. This does not happen because it is technically impossible to send long messages. There usually exists an upper limit on message length, but in practice it is seldom reached nowadays (Murray (1991, 18–19), however, describes a system with a limit of 80 characters and the consequences this limit had on users' behavior). Instead, for one reason or another, speakers opt for short utterances (although in another environment the same speakers can produce longer utterances, see section 2.4), judging them to be more suitable for IM-mediated conversation. This is a perceived affordance.

An example of how the notion of perceived affordances can be applied to CMC research comes in a study by Lee (2007), who analyzes how the ecological factors of IM-communication ("perceived expressiveness of the language, perceived functions of IM, user familiarity with the language, user identification with the language, technical constraints of inputting methods, speed, and perceived practicality of the writing system" (Lee 2007, 223)) influence the perception of affordances, and how the perceived affordances guide the text-making practices chosen by IM users.

Some authors choose to speak almost exclusively of constraints, while their approach is essentially not too different to the approach of those who prefer "affordances". For instance, Anis (2007, 91–96) offers an elegant multiple-constraint-based model of mobile communication.

I use both notions, affordances and constraints (affordances being the hyperonym), and speak about primary (*real*) affordances and secondary (*perceived*) ones. The inability to read a message typed but unsent by an interlocutor is primary: nothing can be done about it. The decision to press Enter at some point while typing a message is driven by secondary affordances. Physically, one can continue, if one wishes (although continuing might not be the best conversation strategy).

Generally, I agree with Hutchby (2001, 86), who, discussing research into telephone interaction, claims that it is best not to look for differences between telephone and FtF conversations that can be attributed to the specific properties of the channel, but to look instead "at how humans have responded to the telephone range of affordances for communication in order to mould specific, and occasionally new, forms of talk-in-interaction". Still, the range of affordances is shaped by the physical properties of the channel, so the role of these properties, if they can be estimated, should not be ignored.

In some cases, however, as mentioned in the previous subsection, it can be difficult to say what is physical and what is social, what is real and what is perceived, and the answer will ultimately depend on whether one takes a realist or a constructivist stance. This higher-order philosophical discussion is beyond the scope of my research. What I am looking at is how the combinations of communication channel properties (that can be viewed as constraints and affordances, primary, secondary, or in some cases borderline) lead to the emergence of new phenomena in language.

Variation and change

The influence of the communication channel on language can manifest itself both from a *synchronic* perspective and from a *diachronic* perspective. In the former case, we are dealing with language *variation*, in the latter with language *change*. Variation and change are, of course, very closely related, and change is the consequence of variation, but a borderline can be drawn. Generally speaking, if we demonstrate that the frequency distribution of linguistic variants is different in different channels, we can hypothesize that the channel has an effect on variation. If we demonstrate that a new variant has come into existence in one of the channels, then we are witnessing language change and have grounds to investigate whether the properties of the channel have played a role in the emergence of innovation.

Most current CMC research focuses on variation, and it is widely believed that no or very little linguistic change has taken place due to the emergence of the Internet (Androutsopoulos 2011a). My focus is predominantly on change, and I attempt to show that changes in linguistic structure that can be associated with the influence of the Internet do exist. In Chapter 3 ("The 'orthographic' neuter: a grammatical innovation in Russian internet language"), a morphological innovation is discussed, in Chapter 4 ("The written turn and its role in actuating language change"), novel forms of interaction between writing and speech, in Chapter 5 (""You were typing something, I interrupted": Play with conversation structure in quasi-synchronous computer-mediated communication"), novel linguistic devices of language play. Chapter 2 ("E-mail vs. chat: the influence of the communication channel on language", discussed in section 2.4), on the other hand, deals with variation, and so does Chapter 6 ("Rise and fall of the Olbanian language: diachronic description of an online slang"), where the diachrony of an online slang is investigated. Whatever the focus of attention, variation or change, the key question is whether the observed phenomena can be linked to the properties of the channel and, if so, in what way.

Direct and indirect influence of the channel

Deutscher (2000, 182), discussing the potential role of writing in the evolution of sentential complementation in Akkadian, proposes differentiating between *indirect* and *direct* correlates of writing, where the former are "the pressures that writing exerts on language through its influence on the structure of society, and thus on communication patterns", and the latter are "the intrinsic linguistic features of written communication, which distinguish it from spoken communication".

This distinction is very relevant to CMC. Schematically, the direct pattern of influence can be described in the following way: affordances of the channel → linguistic behavior of speakers → language variation → language change. The indirect pattern includes one more step: affordances of the channel → *non-linguistic*

behavior of speakers → linguistic behavior of speakers → language variation → language change. Examples of relevant non-linguistic behavior include the choice of topic of the communication, or even the decision whether or not to use a channel (for example, younger people use IMs more often than older people, which means the IM population is on average younger, which obviously affects language variation).

Deutscher (2001, 182–184) is cautious in his conclusions about the role of writing in the evolution of complementation. He is confident of the huge indirect influence of writing, since, inter alia, it has allowed society to achieve a higher level of complexity, which, in turn, has led to increased complexity in communication needs. He finds it difficult, however, to estimate the direct influence of writing, and doubts it could have been substantial.

The innovations I investigate are usually best explained by a combination of the two types of patterns, both direct and indirect. I pay attention to both, although the focus is predominantly on direct influence.

Metaphors of effect and the effect of metaphors

Concluding his review of conceptions and evidence of language change linked to digital media, Androutsopoulos (2011a, 156–157) states:

> Metaphors of ‘effect‘ and ‘influence’ have been common in both [popular and academic] discourses, and the aim ‘to understand the way CMC might affect our language’ (Smyk-Bhattacharjee 2006: 69) has been a legitimate scholarly approach. Alternative metaphors may help us move beyond the implicit technological determinism that still shapes much thinking on language and new technologies (Squires 2010; Thurlow 2007; Androutsopoulos 2006). Such an alternative might be a view of digital media not as containers that determine the language they contain, but as resources for social practices, which do constrain, but do not determine the shapes and styles of network writing.

This extract, together with what I said above, might create the impression that the discussion is solely about which metaphors are more appropriate. Indeed, it is

difficult to disagree with the proposition that the description offered by Androutsopoulos at the end of the extract quoted above does a better job of capturing what happens in reality than a plain "metaphor of effect". It is equally difficult to disagree with the assertion that to *constrain* is still to exert *influence*, so for simplicity's sake it is legitimate to speak of effect and influence, understanding that these are labels for a complex and multi-faceted process.

Unfortunately, there is more behind these two metaphors than a mere choice of words. Those who stick to the "channel influences language" metaphor and to the spirit of the "first-wave" (in Androutsopoulos's terms) CMC studies, are sometimes unaware of sound criticism of many aspects of this approach. A reasonable consequence of this situation would be the claim that the methods used for the study of channel influence should be improved. However, the position that critics of the effect metaphor actually take often seems to be much closer to the assertion that "channel influence is not really worthy of study".

Any study mentioning the i-word can be labeled as suffering from "technological determinism", which, according to Hutchby (2001, 14), is "easily identifiable as the *bête noire* of recent developments in science and technology studies". The label is used so often that its meaning may already have eroded — what exactly is technological determinism in relation to language variation and change? An example of a radical deterministic approach to language variation can be seen in the following quote: "When all the features of the thing described have been taken into account and when the audience has been specified, the form of response is *determined*" (Skinner 1957, 175, emphasis mine). It is unlikely, however, that the Skinnerian approach, famously criticized by Chomsky (1959), would have received much support today. Note that even Skinner talks about taking into account "all the features" (the question of whether it would have helped and whether it is possible remains open), not a single one, like the communication channel.

Rather than believing that technology fully *determines* the linguistic properties of communication, most linguists would agree that it is one of numerous factors that

make certain features more or less probable. So, if a label has to be chosen, something along the lines of "technological probabilism" would suit better.

Even the critics acknowledge that many questions of channel influence are still valid and open (Georgakopoulou 2006, 549). In my view, the most productive approach would be not to evade these questions or replace them with other questions, but to readdress them in a more rigorous way. True, the channel is not the single determining factor, but it is one of the factors, so its influence still has to be studied, and thus the metaphor of influence is still valid.

This, in turn, can be labeled as reductionism, and it is reasonably to ask why a complex phenomenon should be reduced to a simpler model if we know that model cannot fully describe that phenomenon. The aim of science, as I understand it, is to provide the simplest possible explanations. True, in some cases simple explanations are inadequate, and researchers are forced to opt for more complex ones, less formal, more intuitive. Often complexification is necessary, but it is not the goal. The goal is to *improve* understanding, *simplifying* it if possible.

Before proceeding to the next section, where I present a more detailed answer about why it can be worth concentrating on one factor, even if there are many, I would like to justify the use of another metaphor, which is also a simplification device. Croft (2000, 4) warns against the "reification" or "hypostatization" of languages, claiming that languages do not change; it is speakers who change their behavior. Andersen (1989; 2006) makes very similar claims, arguing against leaving speakers out of the picture (this unity is remarkable, since Andersen severely criticized Croft's evolutionary approach to language change). Andersen (1989) also denounces formulations such as "X changes into Y", claiming they are not an adequate model of the reality. Still, in my view formulations such as "feature X changes into/is replaced by/develops into Y" can be used, as long as we remember they are metaphors for 'speakers stop using feature X and start using feature Y in the same contexts". Likewise, a model where innovations *emerge*, *compete, diffuse*, *succeed* or *die out* instead of *being created*, *chosen*, *picked up*, *established* or *abandoned* can be used,

while innovations themselves, of course, do not perform any actions or have any intentions (and are just abstract concepts existing in the researcher's mind).

2.3. Communication channel as a factor in language variation and change

Linguistic structure as an emergent system

The general theoretical view on language that I adopt can be best described as "language is a complex adaptive system". The key features of this system, according to the "Five Graces Group" (2009, 1) position paper, are as follows:

> The system consists of multiple agents (the speakers in the speech community) interacting with one another. The system is adaptive; that is, speakers' behavior is based on their past interactions, and current and past interactions together feed forward into future behavior. A speaker's behavior is the consequence of competing factors ranging from perceptual constraints to social motivations. The structures of language emerge from interrelated patterns of experience, social interaction, and cognitive mechanisms.

In recent decades, there has been growing interest in these competing factors and their role in shaping linguistic structure. Consider, for instance, numerous recent studies of the correlation between social structure and linguistic structure, particularly in the field of language complexity (see Berdičevskij 2012 for overview).

In other words, there is plenty of evidence that language is not an autonomous abstract structure, bestowed upon humans in its entirety, but an ever-changing system whose properties are shaped by evolution, or rather by evolutions: biological and cultural, the latter including changes in language itself, changes in society and changes in technology as well.

Thus, I include the communication channel and its properties among the factors that shape speakers' behavior, and thus linguistic structure. Androutsopoulos (2006, 421) pronounces and welcomes "a shift of focus from medium-related to user-related patterns of language use" in studies of CMC. My approach, as described above,

however, is also user-related, since it focuses on how the properties of the medium affect the behavior of users. The research is about humans, not about technology, and "how technology affects humans" is a shortcut for "what humans do when faced with the constraints and affordances created by technology".

Is this approach reductionist? It does reduce the analysis of a multifactorial problem to one factor, but in no way does it deny the existence or importance of other factors. On the other hand, linguistic structure is considered to be *emergent*, that is, to be a complex pattern that results from a number of lower-level interactions and processes. Exactly how this pattern is formed is not very well understood at present, and it is difficult (if not impossible) to understand fully. To take one step towards better understanding, I choose a simpler sub-problem. In the next section, I discuss how focusing on one factor can be (and has been) productive in science.

Studying a single factor in multifactorial systems

Georgakopoulou (2006, 549) states that "singular interpretations of linguistic choices on the basis of one factor alone do not really work". That must be true, but that does not mean one cannot study the *contribution* of one factor to linguistic choices. If the fact that we live in a complex world where almost nothing is determined by a single factor had stopped researchers from reducing complex problems to simpler ones, our knowledge would have never been advanced.

Consider one of the common ways of assessing the therapeutic and side effects of a medication using a randomized controlled trial, where one group of patients receives the medication in question, while a second group receives a placebo, and then health changes in the two groups are compared. While efforts are usually made to ensure that the two groups are as similar as possible, there are always differences, and the medication is still not the *only* factor affecting their health. Nevertheless, it is considered possible to assess its contribution precisely enough.

When it comes to matters as serious as human life or death, even the factors that affect health in a less direct way attract serious attention, even if powerful methods like controlled trials are unavailable. For instance, Öberg et al. (2011) set themselves a goal as ambitious as estimating the number of deaths worldwide that result from exposure to second-hand smoke. Their answer of 603 000 deaths in 2004 (1% of total mortality) was considered reliable enough to be published in *The Lancet*, one of the most prestigious medical journals.

More subtle influences can be investigated as well. Bushmann and Huesmann (2010) discuss how the effects of televised violence on aggression can be measured. They state that if a *true experiment* (a controlled randomized trial) shows that "children who are randomly selected and shown a more violent show behave more aggressively than other randomly selected children who are treated the same except for seeing the violent show, *it must be that viewing the show caused the children to behave more aggressively.* There is no other possibility" (Bushmann and Huesmann 2010, 230). True experiments of this kind are, of course, hard to perform, so Bushmann and Huesmann discuss other types of studies, both the usual *cross-sectional field studies*, where the researcher collects data by observation and then looks for a correlation between TV viewing habits and levels of aggression, and *longitudinal field studies*, where observation is performed over several distinct time periods. When combined, these two types of studies can provide fairly solid evidence about whether the link between TV violence and aggression exists, and whether it is causal, and, if so, in which direction the causation works. Ideally, however, the data from all three types of studies should be combined (Bushmann and Huesmann 2010, 231).

Returning to the vicinity of the Internet, consider the study by Sparrow et al. (2011). The authors run a series of ingenious experiments to test how the availability of information "at our fingertips" (that is, the possibility to Google nearly anything) affects human cognitive abilities, in particular, memory, and report that there is some effect.

Returning to language, consider the example of dialect studies, or, more broadly, the approach of variationist sociolinguistics on the whole, where linguistic variables are linked to various parameters, such as place of origin, age, gender, social class etc. Androutsopoulos (2011b) demonstrates some limitations in the application of the variationist approach to CMC (see more in section 2.4), but, on the other hand, acknowledges that, according to the accumulated data, "language variation online is patterned by age, gender and region" (Androutsopoulos 2006, 425).

If we consider age, gender and region as parameters worthy of research, why then shy away from channel? Dahlberg (2004), warning against singular causation fallacy, notes that "[e]ven those who prescribe complex levels of determination, such as critical political economists, in the end tend to favor just one determination." Indeed, this is something that should be avoided, but dispreferring some of the parameters is also a fallacy, especially when evidence exists that they can be interesting and important.

In the case of the channel, such evidence exists.

Communication channel and language

In linguistics, the best-researched opposition between communication channels is, of course, oral speech vs. written speech. From the enormous body of work that contraposes the two channels in one way or another, let me mention the classic works of Chafe (1982) and Biber (1988), important analysis by Olson (1996), a detailed review of orality-and-literacy studies by Gee (1986) and a brief overview of the functions and consequences of writing by Coulmas in the introduction to his own book (Coulmas 1989).

When it comes to direct investigation of the role of writing in diachronic *change* of linguistic structure, fewer studies can be named. Still, Wray and Grace (2007, 557–559) outline a number of fascinating possibilities in terms of how writing could have

influenced the development of language. Karlsson (2009) discusses the role of writing in the emergence of finite clausal subordination and concludes that it is significant, since written language relaxed the limit on clausal embedding complexity and made elaborate clausal subordination possible.

Another distinctive channel that has the potential to affect linguistic structure is hand-signing, which, of course, has also been investigated from this point of view (see, for instance, Sandler and Lillo-Martin 2006).

The physical features of the environment also seem to be able to shape linguistic structure. For instance, Fought et al. (2004) find that languages spoken in warmer climates tend to be more sonorant. The reason for this, in their opinion, is that speakers in warm climates on average communicate across larger distances than inhabitants of cold areas, and thus more sonorant languages constitute a communicative advantage. Trudgill (2011, xxxii–xxxvii) lists more examples. Consider, for instance, a complex system of direction prefixes in a Tibeto-Burman language, Qiang (LaPolla 2003, 124), which includes prefixes meaning ‘up-river’, ‘down-river’, ‘up the mountain’, down the mountain’. Unsurprisingly, the Qiang people live on mountain slopes in river valleys. Similar phenomena can be observed in other Tibeto-Burman languages, as well in other language families. Another example from the domain of phonology concerns Australian languages, whose phonological systems are rather unusual compared to other languages. Butcher (2006) notes that, due to chronic middle-ear infections, as many as 70% of Aboriginal children may have significant hearing loss. This loss, however, affects different parts of the frequency scale to different extents, and Australian phonological systems tend to be rich in sounds, the perception of which is likely to be affected *least.*

If these structural features were indeed shaped by the properties of the physical environment, then this was probably a gradual evolutionary process, possibly taking as long as thousands of years. That does not mean, however, that shorter-term processes do not deserve attention. Human-made channels (other than the writing discussed above) can also be expected to influence language to a certain extent:

consider, for instance, the metametadiscussion (sic) of the metadiscussions of the alleged impact of the telegraph on language, in a blog post by Liberman (2011).

Still, of human-made channels writing is probably the most important, and without doubt the best-researched channel – or at least it has been until now. However, research into "how writing changed language" faces an obvious and crucial problem, namely the absence of data reflecting how things were *before* writing. Modern spoken language is not a reliable source, since it is probable that it has been influenced by writing and, more generally, literacy. The same even applies to unwritten languages if their speakers are bilingual and have command of another language which has a written tradition. Data from truly oral societies, those where *no one* can write, should be more important, but are scarce and sometimes contradictory.

The same problem, although to a lesser extent, might be faced in future (or maybe already is now) by those who focus on the role of CMC. First, language use in other channels may be affected by the "language of CMC", with the result that researchers will no longer be able to find an informant capable of producing truly pre-CMC language. Second, as everything becomes digitized and goes online, and the online-offline distinction is progressively blurred (see section 2.4), it may happen that accessing samples of truly pre-CMC writing (let alone speech) will not be straightforward. It seems fairly safe to expect that, in 20 years' time, it will be significantly more difficult to compare "online language" and "offline language" than it is now.

Thus, it is better to do that, or rather to continue doing that, now.

CMC and language

Androutsopoulos (2006, 420) states that we now have "a fairly good understanding" of the "language of CMC", which includes "a wealth of descriptive accounts" of the

"unique features" of CMC, [4] the hybrid combination of written and spoken features, and differences between synchronous and asynchronous registers. This does not, however, seem to be a universally held view. Jones and Schieffelin (2009, 80) claim that "[t]o date, research on CMC has generated more questions than answers, specifically regarding the impact of these technologies on language and language use". Thurlow (2006) observes that "CMC scholarship has tended to focus on the dynamics of interpersonal and group communication rather than on the specifics of linguistic practice".

Thus, there is still much to learn, even though the linguistic properties of the "language of CMC" have already attracted a fair amount of attention. Some studies have already been mentioned above, and I ought to mention some more. The list is not exhaustive, and displays a bias towards studies which deal with IM (which receives more attention in the dissertation than other registers of CMC) and towards those which could be considered to be suffering from a "medium-centered approach", technological determinism or a related vice.

Varnhagen et al. (2010) propose a taxonomy of new language use in IM. Tagliamonte and Denis (2008) compare linguistic variables in two large corpora (IM and oral speech among *the same* teenagers) . Collote and Belmore (1996) and Yates (1996) also compare different corpora, although on a smaller scale and without the advantage of having samples in different channels produced by the same users. Comparison of corpora (e-mail vs. "pen-and-paper" letters) also forms part of Hård af Segerstad's (2002) work, which also includes analysis of log data from IM, chats and SMS, as well as questionnaires and interviews. Log data from IM are also analyzed by Baron (2004), who focuses on gender divergences. Zitzen and Stein (2004) use log data from chats to study how and to what extent the pragmatic, social and

[4] Androutsopoulos uses quotations marks when speaking about the "language of CMC" and its "unique features". I agree with the first use, since, as I mentioned in section 1, it is useful to remember that CMC consists of many registers that are highly heterogeneous (actually the "languages of CMC"), but not with the second one, since some features that some of these registers share can indeed be said to have no equivalents elsewhere.

discourse properties of chat conversation can be explained by "the physical and technical circumstances, and the co-presence conditions following from them, of the Internet chat" (Zitzen and Stein 2004, 983). Werry (1996) provides an analysis of the discursive properties of IRC, arguing that the emerging conventions are "a direct reflection of the physical constraints" (Werry 1996, 48). Condon and Čech (2010) set up an experiment that compares the discourse strategies of subjects engaged in decision-making interactions in three modalities: face-to-face, computer-mediated asynchronous (e-mail) and computer-mediated synchronous. Rowe (2011) shows how e-mail accelerated language change in a private code used by two siblings.

Vandergriff and Fuchs (2009) compare language-play in CMC and FtF discussions among people learning German. Palfreyman and al Khalil (2007) study how the UAE vernacular of Arabic, "under the combined pressure of technical and social change", starts being frequently used in written form (that is, in CMC), and how it is represented in IM. These studies, together with Anis's (2007) aforementioned work and other texts from the same volume (Danet and Herring 2007), can serve as examples of a shift towards non-English-focused CMC research. Further examples include Jarbou and al-Share's (2012) analysis of how male and female speakers of different dialects represent the consonants of Spoken Jordanian Arabic in chat, and Yang's (2007) study of adaptations of the Chinese writing system to CMC. Fung and Carter (2007) focus on English, but on the "hybrid variety" used in ICQ by Cantonese speakers. Gao (2006), in turn, focuses on the "new variety" of Chinese online and the impact the English language has on it.

Herring (1999) analyzes the means of maintaining interactional coherence in quasi-synchronous CMC and, importantly for the purposes of this section, notes that "[t]echnological determinants can not be dismissed as irrelevant in the case of interactional coherence—they have clear effects". Squires (2007) studies variation in IM from a variationist perspective, linking the use of apostrophes and gender, as well as discussing how applicable the concept of standard language is to CMC.

Research is by no means limited to text-only CMC. Keating and Mirus (2003) look into the relationships between technological innovations and new communicative practices based on the example of video communication between deaf users. Naper (2011) investigates how speakers use the affordances of a 3D virtual environment to manage addressivity. Multimodality can be not only the object, but also the instrument of research, see studies by Marcoccia, Atifi and Gauducheau (2008), Beißwenger (2008), Smith and Gorsuch (2004) and Garcia and Jacobs (1999), where the authors do not confine themselves to text-only data, but also make video recordings of a subject's behavior during IM conversations and analyze that behavior.

These studies substantiate my claim that there are reasons to believe that increased attention to the channel does not necessarily signify an illegitimate or unproductive scholarly approach. In the next section, I will review some more specific points of criticism and discuss whether they apply to my own work.

2.4. "Influence of channel" revisited: problems and solutions

The problem of singling out one factor

A recurring critique of the study of the channel's influence is that it is impossible to single out this influence and differentiate it from other parameters. For instance, while it can be tempting to attribute the general informality of IM communication to some inherent properties of the channel, it could instead be to do with the fact that IM users on average are relatively young, and younger people tend to use informal language.

Even when language production among the *same* users across different channels is compared, as in Tagliamonte and Denis (2008), some researchers can still question whether any observed linguistic differences should indeed be attributed solely to the change of channel and not to some accompanying changes, for instance the place of communication (home for IM or school for FtF). Besides, in this specific study,

recordings of FtF communication could have been affected to some extent by the observers' paradox, since the participants were aware they were being recorded.

Controlled studies (like Vandergriff and Fuchs 2009) allow to approach the situation where all but one parameters are the same even closer, but cannot be said to investigate fully natural communication.

In order to investigate whether a change in one parameter — the channel — can trigger changes in linguistic variation, I ran a comparative study (see Chapter 2 of the current dissertation) where nothing but the channel is different, but the data come from natural communication. As a data source, I am using the contents of my own Gmail mailbox. Gmail provides not only the usual e-mail communication, but also a chat system (Gmail chat). Since the chat is integrated into the same window and is easy to use, it is becoming increasingly popular. I collected my chat and e-mail conversations with three people from my contact list. In order to avoid any manifestation of the observer's paradox, I am only using conversations which took place after June 2007 (after I graduated) and before March 2009 (before I submitted a proposal for my current PhD position), that is, when I was neither studying nor working as a linguist and did not have any ideas about studying language and CMC. Thus, all the parameters except the communication channel itself are controlled for. The interlocutors are always the same, and the conversations occur within the same Gmail webpage, but sometimes in the "e-mail" page section, and sometimes in the "chat" section.

I used ten variables to compare the corpora, and five turned out to be significantly different, with the difference being large enough to be considered important. Mean utterance length was lower in chat, the end of sentence was more often left unmarked in chat (that is, not marked by a period or any other sign), as was the beginning of sentence (that is, no capital letter was used to mark it), there were less complex sentences in chat, and more emoticons.

Thus, the change of channel triggers a difference in the choice of syntactic structures, discourse and pragmatic devices, as well as in the level of adherence to norms. These differences are predictable and are generally in line with previous findings relating to e-mail and chat, and their emergence can be quite convincingly explained by means of the primary and secondary affordances of the channels (see the original study for more details).

Androutsopoulos (2012, p.c.) suggested that one of the alternative reasons for the observed differences might be the distribution of topics. Indeed, it is known that users might show some preferences for what is discussed in which channel. Lancaster et al. (2007) found that while college students preferred IM to e-mail, this only applied to communication connected to personal and social relationships. If a similar pattern can be observed in my data, that would still mean that the channel influences language, albeit in an *indirect* way (see section 2.4). In this case, it is possible to test whether this is a plausible hypothesis.

To test the hypothesis, I counted the distribution of topics within each dyad. Each utterance was coded as either *work*, *personal* or *metacommunication*. The latter label was used for messages discussing the choice or use of the channel ("send me the file", "the chat seems to be glitching", "call me" etc.), and for greetings and farewells (since they perform the metafunction of signaling the beginning and the end of the conversation). The number of words in each category was then counted, and divided by the total number of words produced by the given dyad in the given channel. The results are presented in Table 1.

As can be seen, the distributions are nearly the same. Obviously, there is no significant difference in any of the cases, nor is there any notable effect size (dyad 1: $\chi^2(2)=0.029$, $p=0.99$, $\varphi_c=0.12$; dyad 2: $\chi^2(2)=0.095$, $p=0.995$, $\varphi_c=0.06$; dyad 3: $\chi^2(2)=0.001$, $p=0.999$, $\varphi_c=0.02$). Thus, the explanation does not seem to work.

	Dyad 1			Dyad 2			Dyad 3		
	Work	Personal	Meta	Work	Personal	Meta	Work	Personal	Meta
Chat	97.0	1.9	1.1	6.3	82.9	10.8	78.3	17.8	3.9
E-mail	96.8	0.0	3.2	9.4	82.4	8.2	78.9	18.1	3.0

Table 1. Distribution of topics within dyads (proportion of total number of words as a percentage).

It might be further argued that if the classification of topics is more granular (for instance, which work issues exactly are discussed in chat, and which by e-mail), then some differences can be found. Alternatively, one might say that some other factor has not been taken into account (time of day? interlocutors' mood?). Criticism of this kind, however, is *unfalsifiable*: it is always possible to find a potential factor and claim it has not been accounted for. Thus, at some point, when the role of the most plausible factors has been excluded, the burden of proof shifts to those who deny the influence of the channel.

For the purposes of the dissertation it would be ideal, of course, to compare a CMC channel and a non-CMC one in the same way, but this seems to be extremely difficult, if not impossible. The pilot study discussed in this section does not directly demonstrate that CMC makes language different, but it demonstrates a more general point: *the communication channel can make language different*.

The problem of sample size

Androutsopoulos (2006, 420) states that "first-wave" CMC studies, studying, for instance the differences between synchronous (chat) and asynchronous (e-mail) communication, often relied on "small or even anecdotal" samples. This criticism,

however, does not apply to studies such as Tagliamonte and Denis (2008), where the samples are huge.

It might be argued that the sample in the study described above is impermissibly small, but it should be noted, however, that the study aims not to describe the general properties of e-mail and chat, but to provide an (pilot) answer to a fundamental question: does channel influence language? I show that there are differences in language variation across two samples. In terms of linguistic units, the samples are large enough to provide reliable results (12 260 words for e-mail, 17 671 for chats). The results are clearly statistically significant, so they are unlikely to be random.

The problem is that all this has been demonstrated among just four human subjects, and that is indeed a weakness. It is possible that while these four subjects unmistakably demonstrate a change in linguistic behavior when they switch between e-mail and chat, this is some kind of psychological aberration, and nobody else does that. This, does not, however seem very probable, since the changes in behavior are easily understandable and fit well with other descriptions of e-mail and chat. Still, of course, the results would have been more convincing if the sample were larger.

In the other chapters the problem of sample size is less applicable. Again, since I do not aim to describe the "language of IM" or the "language of teenage bloggers", but only aim to describe and explain several linguistic innovations, and in some cases to trace their diachrony, I do not have to collect samples that would be representative of the population of IM users or the population of teenage bloggers.

Probably the most common approach to data sampling in CMC studies is to take a communication log for a certain group in a certain subchannel and analyze it. This, however, is not the best way if one is interested in reviewing as many instances of a particular innovation as possible (although it could be useful for estimating the frequency of the innovation in actual usage).

When I argue that an innovation *exists*, sampling is of little importance: it is enough to provide examples. When I argue that an innovation is frequent, or compare it with

an alternative variant (Chapter 3), or trace changes in frequency (Chapter 6), the samples I use are quite large, both in terms of linguistic units and language speakers (a large forum or another participatory website, Russian blogosphere on the whole etc.).

The problem of the non-ethnographic approach

Androutsopoulos (2008) describes the limitations of approaches based exclusively on log data, assuming that they do not make it possible to examine the discourse practices of the participants and relate them "to observable patterns of language use", and suggesting that they should be complemented by more ethnographically-oriented studies. In another overview, Androutsopoulos (2011b) discusses the limitations of quantitative methods (specifically, the quantitative premise of variationist sociolinguistics), since they fail to take into account infrequent features, even though they may be important, and proposes not to "shy away from the importance of singular, unrepeated instances of linguistic difference if and as used in a strategic, yet non-quantifiable way" (Androutsopoulos 2011b, 280).

Generally, these two points are very fair, as long as it is remembered that ethnography and qualitative examinations should *complement* analysis of log data and quantitative methods, not *replace* them. Which methods should be given priority depends on the particular research question. In my quest for linguistic innovations and the combinations of factors that triggered them, I give priority to what Androutsopoulos calls analysis of "log data". Interestingly enough, if one takes all the available context into account, one can often infer quite a lot of information that is of ethnographic interest. Consider Chapter 3, where I qualitatively analyze various examples of occurrence of a morphosyntactic innovation, including metalinguistic discussions about it or its use in language play, in order to argue that it has evolved from a rare playful device with specific connotations into a popular and productive grammatical phenomenon without obvious semantic load.

As for quantitative methods, they are a powerful tool, indispensable when it comes to testing predictions that follow from hypotheses, but in some cases their applicability is indeed limited. Their use varies from chapter to chapter. Chapter 2 and Chapter 6 are both fully based on quantitative methods, which makes it possible to describe the changes in popularity of an online slang over 11 years, both within the community of its "core" users and among the wider public. In Chapter 3, conclusions from qualitative analysis are supported by quantitative evidence. In Chapter 4, no numerical data are provided, since the chapter is meant to be an overview of a trend, not a detailed description, but further investigation in some of the research directions outlined in the chapter requires quantitative verification. Chapter 5, however, is fully qualitative, since quantitative methods are not applicable to the data, at least not in any straightforward way.

The problem of homogenization

Androutsopoulos (2006, 420) criticizes "first-wave" CMC studies for implicitly assuming that language use in specific channels is relatively uniform and homogenized. He notes that "[i]t is empirically questionable whether in fact anything like a 'language of e-mails' exists, simply because the vast diversity of settings and purposes of e-mail use outweigh any common linguistic features".

It is indeed empirically questionable, and it is not obvious that the answer is "no". True, the linguistic properties of e-mails can differ tremendously depending on other parameters of communication, but still it is possible that a description of "prototypical" and "peripheral" e-mail styles or distinctive e-mail features can be produced. It will necessarily be simplified and overhomogenized to some extent, but virtually any description of any idiom is. The vast diversity of settings, purposes and competences can also outweigh the common linguistic features of a sociolect, a dialect or a language. This, however, is not the question that I intend to answer, and, in my case, the problem of homogenization is not acute. I am making minor

generalizations about language use in particular channels. Most frequently, I simply claim that some linguistic pattern exists, and often discuss when it came into existence and why.

Where possible, I present and discuss data in such a way as to minimize homogenization. In Chapter 6, for instance, a box-and-whiskers plot is used to present graphically diachronic changes in the frequency of certain items. While this type of plot might seem complicated to a reader who is unfamiliar with it, it makes it possible to visually represent many properties of the distribution of a variable (median, interquartile range, outliers etc.) and not only the arithmetic mean, as is sometimes the case in the humanities.

The problem of establishing causality

Another weakness of "channel influence" studies relates to claims about causality. How can statements that a channel affects, triggers or causes something be supported with evidence: in other words, how do we establish that there is indeed a causal link?

First, the most obvious criterion is the cross-channel comparison, as discussed above (3.4, 4.1). If a linguistic feature is present in one channel and absent in another, that is an indication that it may have something to do with the channel. The main problem here is that it can be quite difficult to prove that a linguistic feature is absent somewhere. The Internet is searchable (not ideally, but reasonably well), while pre-CMC channels are less so. Still, there exist corpora of spoken and written speech, and digitized archives of press and books, so at least some data are available.

The problem is less acute with lexical units, since they are easy to search for, and more acute with something less concrete and "tangible", like structural features. Some solutions, however, can usually be found. In Chapter 3, I devise a method for estimating the frequency of a morphosyntactic agreement pattern, in Chapter 6, of antinormative spellings. It is important to realize, of course, that the results are

always approximations of the reality, but if different approximations show a consistent pattern, they should be taken seriously. In other words, the choice is either to search as thoroughly as one can and then take absence of evidence as evidence of absence, or to remain agnostic. I opt for the former.

Second, the observed differences between channels have to be accompanied by a theoretical explanation. Why do they exist? What is the mechanism of their emergence? Sometimes, the suggested explanatory theory makes it possible to formulate predictions that can subsequently be tested in order either to support or refute it.

As was mentioned above, quantitative studies represent one tool for testing predictions. In Chapter 3, for instance, the prediction is that the choice of an innovative agreement pattern is to a large extent determined by the formal properties of the noun. Comparing how frequently different agreement patterns are used with different nouns, I show that this turns out to be true.

Another way of verifying a researcher's interpretation is to compare it with lay speakers' intuition, either by analyzing metalinguistic discussions in CMC or by conducting interviews and surveys. Any explanations of linguistic phenomena gathered in this way should, of course, be viewed with great caution, but if they differ significantly from what the linguist hypothesizes, then the difference should be explained.

Experiments might be another way of testing some claims and assumptions. In Chapter 6, for instance, I use the assumption that a text rich in antinormative spellings is difficult to read (and to write as well). This view seems reasonable, is shared by many linguists, makes it possible to explain the data well, and is supported by metalinguistic observations of the speakers, but can in principle be verified even further by a direct psycholinguistic experiment.

In this dissertation, the evidence is largely limited to demonstrating difference between channels, explaining it and, where possible, providing some additional

support to the explanatory hypotheses. It is important, however, that these hypotheses can be tested and developed further, and this section outlines some ways of how this could be done.

The problem of differentiating online and offline

Finally, I would like to discuss one more methodological issue that, to the best of my knowledge, has received little attention in the literature, but still deserves to be kept in mind. I am grateful to Roman Leibov for bringing it to my attention at the "F4" conference of the "Future of Russian" project.

When we compare CMC and "pre-CMC" written language, what exactly do we mean by the latter? Documents, letters, notes, shopping lists and home assignments at school are increasingly typed on a computer (and then printed out, or maybe even not). Books are now also drafted using a computer, and often read on a computer, and published online, and sometimes even prepared *exclusively* for use online or on an iPad screen. A newspaper article — can it be considered an example of an offline text, even if it is printed on paper? It was typed using a keyboard, formatted on a computer, was most probably sent to and from editors, copy-editors, designers and fact-checkers by e-mail or another means of exchanging files, and is most probably published online as well, where it can be commented on by website visitors.

That said, there are still obvious differences in terms of primary and secondary affordances between a forum, a newspaper that is published online only, the website of a paper newspaper, a paper edition of this newspaper and a printed book. Intuitively, I would say that they form a cline from "most online" to "most offline", and are listed here in decreasing likelihood of adopting the innovations that I study. My primary interest lies in the registers where the innovations emerged and managed to survive their early days (which can be difficult for a new feature in language), and that seems to be the online side of the cline, with its unique combination of physical affordances and their social use.

2.5. Conclusions

In this section, I have addressed a methodological paradox: while some CMC scholars speak of the "undeniable influence" of technology on language (Thurlow 2006, 668) and defend doctoral dissertations with titles like "The Influence of Computers, the Internet and Computer-Mediated Communication on Everyday English" (Greiffenstern 2009), others propose abandoning the very metaphor of influence and denouncing the approach it epitomizes.

It should be noted that the critical reaction of Androutsopoulos and others was to a large extent triggered by early publications on language and CMC, as well as by numerous publications in the media, non-academic but influential in terms of affecting public opinion. Some of these works were indeed unreasonably deterministic, as well as suffering from other drawbacks. Thus, the reaction is entirely understandable, but it is important that it does not develop into overreaction.

In this section, I have attempted to offer some kind of synthesis, taking into account the important points in the criticism, but adhering to the basic research questions of the "channel influence" approach, and arguing that the metaphor is valid and convenient, and can be used productively. This is what I aim to do in my dissertation.

As mentioned above, my methods, for the most part, do not produce exact answers or absolutely conclusive evidence (this does not often happen in linguistics). They, do, however, make it possible to accumulate certain facts, suggest explanations and argue in favor of these explanations. It is up to readers to estimate how convincing my conclusions (presented in section 4) are, but they certainly can be verified and refined (or refuted) if necessary.

3. More about framework

In section 2.3, I have already addressed the question of my theoretical stance, but I would like to revisit it for clarity's sake. Linguists differ in their views on which role frameworks should play in research. Some believe that no linguistic research is possible unless it is based on some underlying theory (Kibrik 1992, Testelec 2001). Some, on the contrary, claim that the most productive way to perform research is to do so framework-free (Itkin 2007, Haspelmath 2010).

Why might one need a framework? In my view, the basic advantages it brings are as follows. First, it provides insignia: to what camp does one belong? Which theoretical assumptions does one have? Second, it provides a methodology and a metalanguage: how does one analyze the phenomena observed and how does one describe them? Third, it provides a roadmap: which research questions should one ask? And how does one interpret the answers?

For insignia, I choose the complex adaptive system (CAS) approach, outlined in section 2.3. It is not explicitly mentioned in the chapters, but I use it here to present and overview the results. Thus, the general model of language I have in mind includes the following features: variation at all levels of language; the probabilistic nature of agents' linguistic behavior; continuous change in the patterns of this behavior throughout the agents' lifetime (i.e. even adult grammars are not fixed); the influence of usage patterns on grammar etc. ("The Five Graces Group" 2009). The CAS approach, however, does not provide a full-fledged metalanguage or roadmap, at least not in the general form of it on which I rely. Thus, while these two aspects of my work are CAS-inspired to some extent, they are actually framework-free.

As regards methodology and metalanguage, they depend on the particular question the chapter addresses. In Chapter 5, I use the conversation analysis approach (which in itself is a bottom-up, data-driven approach with rather simple and transparent metalanguage, and without many complex theoretical assumptions). In Chapter 4, I use the metalanguage of "classic" variationist sociolinguistics, although the roadmaps

and methodologies are quite different. The approach in Chapter 2 is typical of corpora-based register studies. In Chapter 3 and Chapter 6, no predefined label fits well, but the methodological solutions are explained where necessary and the metalanguage is, in my view, transparent (and potentially translatable into other metalanguages).

As regards roadmaps, I outlined and justified the basic questions in the previous section, and more specific questions are formulated in the chapters. The absence of a full-fledged roadmap can sometimes be felt. In Chapter 6, for instance, the diachronic description of an online slang does not make it possible to answer the question of whether there are important differences in the patterns of development of this slang and other online and offline substandard varieties, since there are no comparable diachronic descriptions. Not much can be done about it, the field is still new, and a first step has to be taken at some point. I believe future research will provide data of comparable nature. When it comes to *proposing* a roadmap, in Chapter 4 I outline a research direction which, in my view, is most promising for the study of linguistic innovations in CMC.

Thus, the dissertation is heterogeneous in terms of approaches and data samples. This, however, is predetermined to some extent by the research questions. The emergence and diffusion of innovations in CMC is a multi-faceted process, and various aspects of it require different approaches. To conclude this discussion, I would like to draw a comparison between the field of language and CMC and the contemporary evolutionary biology, as described by Markov (2010, 19–20, translation mine): “Contemporary biology is not even a patchwork quilt, but a rapidly growing pile of patches, and one can only try to guess how the prospective ‘quilt’ will look like, and one never knows in advance which guess was correct and which one will require retailoring.”

4. The study of language change online: results

As I mentioned in section 2.4, in Chapter 2 I study linguistic variation across two different communication channels: e-mail and chat. The results show that out of ten variables, five do not depend on the channel parameter: mean word length, lexical density (the ratio of lexical items to the total number of words), type/token ratio (the ratio of *different* words to the *total* number of words), the frequency of first-person pronouns and the frequency of brackets. All variables, apart from the last one, have been shown to be typical for *informational* speech production (more typical for "written-like" registers), as opposed to *involved* (more typical for "spoken-like" registers), and it is reasonable to expect that the intensity of bracket use would demonstrate the same pattern.

Two other variables — mean utterance length and the frequency of complex sentences — which are also often used to estimate a place of the register on the informational–involved scale, demonstrate significant differences. Utterances are shorter in chat, and complex sentences are used less frequently. In addition, speakers use more emoticons in chat and more frequently ignore two rules of punctuation: mark the start of a sentence with a capital letter and mark the end of it with a period (unless it is a question or an exclamation).

How can we interpret these differences from the CAS approach point of view? Users adapt their behavior to the constraints and affordances of the channel. Chat, with its narrower window, higher speed of communication and lower stability, pressures users into shorter and simpler utterances, which explains the syntactic differences. Chat is also perceived by speakers as a more informal channel, where communication is more involved, more personal, more phatic. Although this does not lead to any differences in the choice of topic (see section 2.4), speakers are more eager to compensate for the lack of non-verbal cues by using more emoticons. It is also possible that emoticons serve as a means to economize on time and words: instead of crafting a sentence to convey intonation unambiguously, it is possible to denote it with an emoticon. The perceived informality of chat, together with the desire to use

its potential for high speed, also leads to some violations of the standard norm: the start and end of a sentence are often left unmarked, something that almost never happens in e-mail. Another important reason for that is that the norm is actually redundant: if a message contains only one sentence (and that is usually the case), then even without capitals and periods it is clear where the sentence begins and where it ends.

Thus, the communication channel plays some role in determining how eager speakers are to use certain patterns of syntactic and pragmatic behavior, and to follow the norms. The linguistic variation is noticeably affected, and variation is a nutritional medium for change.

In Chapter 3, I move from variation to change and discuss the role of CMC in ensuring the diffusion of a morphosyntactic innovation in Russian, namely, the playful conversion of nouns of masculine and feminine genders to neuter: *краси́вое дéвушко / krasívoe dévuško* 'beautiful-NOM.SG.N girl-NOM.SG(N)'[5] instead of *краси́вая дéвушка / krasívaâ dévuška* 'beautiful-NOM.SG.N girl-NOM.SG(F)'; *моё́ му́жо / moë múžo* 'my-NOM.SG.N husband-NOM.SG(N)' instead of *мой муж / moj muž* 'my-NOM.SG.N husband-NOM.SG(M)'.

Such a conversion was occasionally used in fiction, poetry and the press throughout the 20th century and even earlier. Two things, however, are important here. First, these were occasional uses. In other words, they were instances of one-speaker innovation. In the early 2000s, however, the innovation was picked up by many more speakers and became a frequent and productive model.

Second, prior to that it was used as a salient artistic device, usually adding certain (often pejorative) connotations to the meaning of a noun (for instance, it could imply the asexuality of a denoted person). Now, on the contrary, having diffused widely

[5] The conventions of The Leipzig Glossing Rules (http://www.eva.mpg.de/lingua/resources/glossing-rules.php) are followed.

enough in CMC, the innovation has lost most of its original semantic connotations, and has become a means of adding some general ironic flavor to one's utterance. This is an important point: linguists often propose specific semantic explanations for the instances of the use of new neuter. In the chapter, I show that in many cases, no plausible meaning can be ascribed to the new neuter, and none of the proposed meanings works well for all the examples. It is used rather for pragmatic purposes: to show a trace of irony in one's speech, not to sound pompous, in some cases, to shed some of the responsibility for a judgment.

How can we explain the sudden success of the innovation in CMC? As I mentioned above, the scope for this had existed in the Russian language system previously, and individual speakers made occasional use of it. Other speakers, however, did not adopt it. In the written medium, the innovation had slim chances of spreading. Public channels were controlled by editors, and while a single deviation from the standard norm could be allowed in fiction and poetry, it had no chance of acquiring any status other than that of a seldom-used artistic device. Private written channels (i.e. personal letters, diaries etc.) did not have external editors, but were a poor medium for the diffusion of innovation since, unlike CMC, they did not form a multiplex network where information could be transmitted rapidly and one-to-many (potentially one-to-all). Moreover, there was no triggering event which could have launched innovation diffusion. Nor was oral speech a perfect medium for the diffusion: due to vowel reduction in unstressed positions and allegro-pronunciation effects, potential new-neuter forms and standard forms are indistinguishable in many cases (*дéвушко / dévuško* is pronounced in exactly the same way as *дéвушка / dévuška*). In other ways, neither were there good conditions for the diffusion, nor was there any triggering event.

Interestingly, it was specifically the phonological neutralization effect that made the triggering event possible. In the early 2000s, an anti-standard idiom called *Olbanian language* became immensely popular in Russian CMC (see Chapter 6). One of its basic principles is to 'break all rules of orthography', and one manifestation of this

principle is the use of *o* instead of the unstressed *a* (which violates orthographic but not phonological norms: in unstressed positions both vowels are reduced and sound the same). Thus, nouns with unstressed ending — *дéвушка*, *мужчи́на* / *dévuška*, *паи́ка*, *mužčína* / 'girl', 'man' — sometimes assume the form *дéвушко*, *мужчи́но* / *dévuško*, *mužčíno*.

Nouns ending in *-o* in Russian are strongly linked to the neuter gender. As a consequence of this link, words like *дéвушко* / *dévuško* (originally feminine) or *мужчи́но* / *mužčíno* (originally masculine) can be reclassified as neuter, with their declension and agreement following suit. This is exactly what happened, with Olbanian serving as the triggering event (later the popularity of this slang faded and the connection between it and the new neuter was lost).

What is important is that the new register now exists, in the form of text-based informal CMC. This register is written, and thus *дéвушко* / *dévuško* does not coincide with *дéвушка* / *dévuška* and can exist and develop. It is not edited, and norm deviations, both conscious and unconscious, can flourish (although remember the discussion in section 1: not just *any* deviations). Finally, it is a multiplex high-speed network where anybody can write and anybody can read and innovations can thus diffuse rapidly.

The new neuter quickly gains popularity, starting as an orthographic phenomenon, but later spreading to other noun classes (cf. *мужо* / *mužo* instead of *муж* / *muž*). Once the innovative pattern becomes established enough, speakers use it to generate some really unusual forms (see chapter) and even use it in oral speech, going as far as violating phonological rules and pronouncing *дéвушк*[o] / *dévušk*[o].

Another important factor in the success of the innovation was its ability to express a subtle pragmatic meaning, relevant to Russian speakers nowadays — slight irony, subdued pomposity. Discussion of the reasons for this trend, visible in fiction, in press and in everyday talk, is beyond the scope of this dissertation. It could be the consequence of some universal trend, or it could be the result of speakers being tired

of Soviet style (pompous and official). It is also possible that CMC, with its high informality and anonymity, also plays a role in promoting this trend towards slightly ironic, not entirely serious communication.

It is easy to see that CMC ensured the success of the innovation not because it provided some affordances the speakers used to *invent* the new neuter (it had been invented previously many times), but because it provided a suitable medium for its *diffusion*. This aspect of CMC's role in language change is discussed in Chapter 4.

Chapter 4 has two aims. First, in it I describe some aspects of the so-called written turn, the trend towards greater autonomy and salience for written speech. Second, I use this description as an example when addressing the well-known *actuation problem* of sociolinguistics, with relation to the role of CMC in actuating language change. In addition, I discuss some limitations in the applicability of traditional sociolinguistic methods to the study of language change in CMC.

I review the following facets of the written turn: first, the greater autonomy of written speech, second, its greater salience, and third, its increased ability to trigger and direct language change.

The first facet is exemplified by a non-standard emphatic letter elongation. Traditionally, in standard written Russian it is possible to express emphasis by mimicking an emphatic device of oral speech, namely, by elongating one or more letters in a word. The letters have to denote sounds that would have been elongated in emphatic pronunciation. In informal CMC, however, it is possible to express emphasis by elongating *any* letters, even if they denote sounds which cannot be elongated in oral speech or do not denote any sound at all. Consider: *пожáлуууууйста* / *požáluuuuujsta* 'please'; *убиииивáть* / *ubiiiiivát'* 'to kill' (a non-stressed vowel is elongated, this is impossible in oral speech) or *вещьььь* / *veŝ'''* '[cool] thing' (the soft sign *ь* / *'* is elongated, which does not denote any sound and in this particular word does not have any phonetic value at all). This novel pattern shows that people can *write* without thinking of how they *speak*, that the visual

channel can become decoupled from the auditory channel. This fact challenges the Bloomfieldian view of writing as a mere "recording" of language (the repercussions of which are still strong among linguists) and supports the view that writing is as natural as spoken language.

The second facet is exemplified by several cases of features being borrowed from written speech into oral. The trend is, of course, not new per se, but the borrowed features are new, and, interestingly, not the most easily borrowable: emoticons *xə* (in oral speech [hd] or [hede]; etymologically stemming from *xД*, which, in turn is a transliteration of *xD,* which is a representation of a laughing face) and *0_o* (in oral speech presumably [o]; etymologically a representation of the eyes of a surprised person). It shows that speakers do not perceive emoticons as features confined to the written space, but are ready to invent a way to pronounce them, thus enriching oral speech with devices borrowed from written speech.

The third facet is exemplified by the new neuter. As I show in Chapter 3 and briefly summarize above, its diffusion could have occurred only in a written channel. The trend towards greater autonomy of writing made the initial use of forms like *дéвушко / dévuško* possible, while the trend towards greater salience enabled their penetration into oral speech.

How are these observations related to the actuation problem? The actuation problem was formulated by Weinreich, Labov and Herzog (1968, 102) as follows: "Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times?"

Discussing the role of CMC in actuating language change, I highlight the importance of distinguishing between *speaker innovations* and *community innovation*, the latter being the result of the *diffusion* of the former across speakers. Further, I claim that most of the influence that CMC exerts on language change occurs at the stage of diffusion. Consider the new neuter: speaker innovations that could have actuated this change appeared many times, but disappeared without trace. Only when the

conditions necessary for successful diffusion were created by CMC did the pattern achieve the status of community innovations.

For virtually any feature which is presumably new and associated with CMC, some older equivalent can be found at the level of speaker innovation (emoticons, as discussed in the chapter, date back to 1887). Few of these speaker innovations, however, are diffused. Their success in diffusing depends on selective forces which operate in a given language at a given time. These selective forces depend on many internal and external factors, including the communication channels which are currently in use in society.

Thus, I propose the following partial solution of the actuation problem: *A particular change X can take place in a particular language at a given time because selective forces which operate in this language at this given time make possible the diffusion of particular speaker innovations, namely, those that can trigger change X.*

CMC, for instance, reinvigorates the written turn, and thus the speaker innovations which can easily diffuse in written speech, but not in oral speech, enjoy much higher chances of success.

The influence of CMC is not, however, limited to the increased role of writing, and, in Chapter 5, I look at other consequences of its use.

The basic question of Chapter 5 is: do the properties of CMC affect how its users play with language? More specifically, how do users play with conversation structure in quasi-synchronous CMC (QS-CMC). Quasi-synchronicity, typical for chats and instant messengers, means that, while messages are delivered instantaneously, the process of message *transmission* is not synchronous with the processes of *production* and *perception*. In other words, the message-in-progress is visible only to its composer, but not their interlocutors (until the “Send” button is pressed). Quite extensive research has been done on the peculiarities of the conversation structure in QS-CMC. I address the role of the constraints and affordances of QS-CMC in shaping the playful conversational behavior of its users.

Why is it important to study how CMC affects language play? First, play is important per se, it is a ubiquitous and prominent aspect of linguistic behavior. If its patterns undergo any changes when speakers start using a new communication channel, the reasons for these changes and their nature are of great interest to linguists. Second, as observed by Sannikov (2002), language play can be used to study the properties of the language itself. Sannikov argues that language play is a linguistic experiment performed by language-speakers. Since there is always some kind of anomaly underlying language play, we can use the data from this ongoing ubiquitous experiment in *anomaly* to study what is *normal.*

In the chapter, I analyze 12 examples of play with conversation structure. In each example, speakers make use of one or more of the following affordances: quasi-synchronicity per se; availability of info messages ("xxx is typing you a message...") and differences in the speed of typing (both inherently related to the quasi-synchronous nature of communication); issues with the encoding of Cyrillic characters; automatic banning of users violating certain rules of conduct; and persistence of transcript. In order to show these affordances can be playfully used by speakers and to show how the analysis of their playful use can enrich our understanding of the conversational organization of QS-CMC, it is best to review an example from the chapter.

ххх: ты что-то печатала, я тебя перебил.

xxx: ty čto-to pečatala, â tebâ perebil

xxx: you were typing something, I interrupted

Apparently, xxx has sent a message to his interlocutor (let us call her yyy) while she was typing a message to him (and this fact could have become known to him only through an info message "yyy is typing you a message"). As follows from xxx's apology, it must be the case that yyy did not finish the message-in-progress; moreover, she must have stopped typing it.

The situation is similar to an *interruption* in an oral face-to-face (FtF) conversation, and xxx makes use of this similarity, offering a mock apology. The comic effect is created for the most part by the fact that the interruption was not real (this *partial similarity* of linguistic practices in CMC and in real life serves as a principal means for creating comic effects in many other examples). The sequence of actions of the interlocutors is the same in a typical FtF interaction, and in this example: Alice starts producing an utterance, Bob produces an utterance while Alice's is yet to be finished, and this causes Alice to stop producing her utterance. The effect of Bob's interrupting utterance in oral speech, however, is different from the one it can have in QS-CMC.

In synchronous oral speech, it is difficult for two people to produce messages simultaneously, due to the physical properties of the communication channel, but in QS-CMC it is perfectly possible. The "interruption" did not make it impossible for yyy to continue message production.

On the other hand, the "interruption" did indeed cause her to stop typing her message. Most probably, she stopped to read it, and then later did not resume typing, deciding either not to finish the message, or to change it, taking into account what xxx had just said, or for some other reason, cf. Beißwenger's (2008, 14–15) analysis of a video recording of two chat participants modifying their work on the message-in-progress (continuing it, or stopping typing, or deleting the text) with respect to the interlocutor's behavior (typing, or being silent, or sending a message which is presumably not a complete utterance etc.).

It has been claimed that simultaneous talk, such as overlaps and interruptions, is impossible in QS-CMC (Garcia and Jacobs 1999, 346). As regards overlaps, the truth value of this statement depends on how *overlap* is understood. Simultaneous production of utterances, as shown by several of my examples, including this one, is possible. However, simultaneous production of utterances that are immediately available to all speakers (as in oral speech) is indeed impossible. As regards interruptions, while physically splitting the utterance-in-progress with another

utterance is, of course, impossible, interruption-like situations, such as in the example, do emerge in QS-CMC.

González-Lloret (2011, 312) states that "there is no competition for the right to submit a message" in QS-CMC, but xxx behaves as if there is such a competition, and he inadvertently deprives yyy of her right to submit a message. We do not know how serious he is. He may just be joking, or he may be completely serious. Most likely, however, he is half-serious: having observed that there exists a pattern in CMC that is similar to an oral speech pattern, but not identical, he, in his desire to amuse his interlocutor, behaves as if it were entirely identical.

This example contradicts Zitzen and Stein's (2004, 993) observation that, in chat, "there are no possibilities for monitoring the temporary suspension of the right to speak". Info messages provide this possibility, and speakers use it (although sometimes in unexpected ways), thus challenging another claim by Zitzen and Stein, namely, that overlaps are the norm in the composing of messages. While it is undoubtedly true that overlaps in producing messages often occur, the example shows that they are not always viewed as a norm, and sometimes users try to avoid them.

Thus, the constraints and affordances of CMC do play a role in shaping speakers' playful behavior, and in shaping "serious" conversational practices as well. There are reasons to hypothesize that conversation structure becomes more "tangible" for the speakers, they are more aware of their ability to shape it, to change it — and to play with it.

Chapter 6 differs somewhat from the other chapters. It does not directly address the question of CMC's role in the emergence of innovations, but focuses on the diachrony of the so-called Olbanian language, also known as *язык падонков* / *âzyk padonkov* 'the language of scumbags'. Olbanian was an important phenomenon in the Russian Internet (and probably the most widely known and hotly debated one), and its popularity lead to the emergence of various innovations (for instance, the

aforementioned new neuter). The basic idea of this slang (or anti-language) was rebellion against literacy and the standard linguistic norm.

While the slang received a fair amount of linguistic attention, most of the conclusions about its development or the use of linguistic devices are based solely on researchers' intuition. I provide an empirical quantitative verification of some common assumptions about Olbanian. There are a number of important results.

First, I confirm that the anti-normative spirit is manifested predominantly at the orthographic level, and the vast majority of the orthographic norm deviations follow the phonetic principle, i.e. "write as you hear". This is unsurprising, since *padonki* was a protest not against language, but against linguistic norms. Orthography, at least in the case of Russian, is the embodiment of the norm: an artificial and rigid system, codified more strictly than any other language domain and quite difficult to learn due to the lack of one-to-one mapping with oral speech. It is potentially interesting to compare these data about *intentional* errors with the actual frequencies of *unintentional* errors (made, for instance, by schoolchildren).

Second, I calculate the frequency of actual usage of the slang's distinctive linguistic features — that is, deviations from the standard norm — from 2001 to 2011, both *inside* the Olbanian community (on a dedicated website) and *outside* it (in the whole of the Russian blogosphere).

The results show that, inside the Olbanian community, there is a sharp decrease in frequency from 2001 to 2002 and then a gradual decrease from 2002 to 2011. This generally corresponds to the existing view of Olbanian history. In the beginning, speakers toyed extensively with new anti-orthographical devices. Producing and perceiving texts that are oversaturated deviations, however, is quite difficult, and after the initial peak of interest, speakers started using them more economically and in a more symbolic way. It is somewhat surprising that this happened so early.

Outside the community, however, the results are completely different. The frequency peaks in 2006 with a gradual increase in preceding years and gradual decrease in

subsequent years. Again, this roughly corresponds to the existing view that public interest in the Olbanian language peaked in 2004–2006. Interestingly, this peak is not reflected inside the community.

Both inside and outside the community, by 2011 speakers had almost entirely abandoned using erratic spellings and forms. This tallies with the common view that Olbanian has nearly died out. However, the popular hypothesis that some erratic spellings were lexicalized and survived is not supported by my data.

As I mentioned above, these data alone do not allow me to draw any reliable conclusions about how slangs develop, or about how online (as opposed to offline) slangs develop, or how Russian slangs develop, since no comparable data are available. I hope, however, that such data will be accumulated in the future, at least for relatively accessible online slangs.

At least one thing, however, can be said about the role of CMC. As follows from what was said above, the Olbanian slang is a predominantly *written* phenomenon, and so it required a written channel to develop. As in the case of the new neuter, some occasional uses of intentional anti-orthographic spellings can be found throughout the 20th century, but it was only when CMC provided a suitable medium that full-fledged slang emerged.

To sum up, language change online is not exactly the same as language change offline. Moreover, language change offline is not exactly the same now as it was twenty years ago, and the spread of CMC has played a role in this "change of language change". I do not intend to overestimate the importance of CMC — as I have said many times in this chapter, it is just one of the factors in a multi-factor complex adaptive system. Still, the mechanisms of its influence, the description of *how* people adapt their linguistic behavior to the constraints and affordances of the new channel, and the explanation for *why* do they do it this way, are important to achieving a better understanding of how language functions and changes.

Acknowledgments

This chapter benefited from discussions with Jannis Androutsopoulos, Jane Stuart-Smith, Lars-Eric Christer Johansson, Ingunn Lunde, Torjus Midtgarden, Rasmus Slaattelid Aleksandr Piperski and Boris Iomdin, to all of whom I am deeply indebted.

References

Andersen, Henning. 1989. "Understanding linguistic innovations". In *Language change: contributions to the study of its causes*, edited by Leiv Egil Breivik and Ernst Håkon Jahr, 5–28. Berlin: Mouton de Gruyter.

Andersen, Henning. 2006. "Synchrony, Diachrony and Evolution." In *Competing Models of Linguistic Change: Evolution and Beyond*, edited by Ole Nedergaard Thomsen, 59–90. Amsterdam, Philadelphia: Benjamins.

Androutsopoulos, Jannis. 2006. "Introduction: Sociolinguistics and computer-mediated communication." *Journal of Sociolinguistics*, 10(4): 419–438.

Androutsopoulos, Jannis. 2008. "Potentials and Limitations of Discourse-Centred Online Ethnography". *Language@Internet 5*. Accessed November 27, 2012. http://www.languageatinternet.org/articles/2008/1610

Androutsopoulos, Jannis. 2011a. "Language change and digital media: A review of conceptions and evidence". In *Standard languages and language standards in a changing Europe*, edited by Kristiansen Tore and Nikolas Coupland, 145–161. Oslo: Novus.

Androutsopoulos, Jannis. 2011b. "From variation to heteroglossia in the study of computer-mediated discourse." In *Digital Discourse: Language in the New Media* edited by Crispin Thurlow and Kristine Mroczek, 277–298. Oxford: OUP.

Anis, Jacques. 2007. “Neography: unconventional spelling in French SMS text messages”. In *The Multilingual Internet*, edited by Brenda Danet and Susan Herring, 87–115. Oxford: OUP.

Baron, Naomi. 2004. “See You Online: Gender Issues in College Student Use of Instant Messaging”. *Journal of Language and Social Psychology* 23: 397–423.

Beißwenger, Michael. 2008. “Situated Chat Analysis as a Window to the User’s Perspective: Aspects of Temporal and Sequential Organization.” *Language@Internet* 5, article 6. http://www.languageatinternet.org/articles/2008/1532, accessed January 23, 2013.

Berdičevskij, Aleksandr. 2012. “Âzykovaâ složnostʹ [Language complexity. A review]. *Voprosy âzykoznaniâ* 5: 101–124.

Biber, Douglas. 1994. “An Analytical Framework for Register Studies.” In *Sociolinguistic perspectives on Register*, edited by Biber Douglas and Edward Finegan, 44–56. New York: OUP.

Biber Douglas. 1988. *Variation across speech and writing*. Cambridge: CUP.

Bushman, Brad and L. Rowell Huesmann. 2012. “Effects of Televised Violence on Aggression.” In *Handbook of Children and the Media*, edited by Dorothy Singer and Jerome Singer, 223–254. London: Sage.

Butcher, Andy. 2006. “Australian Aboriginal Languages: Consonant-Salient Phonologies and the ‘Place-of-Articulation Imperative’.” In *Speech Production: Models, Phonetic Processes, and Techniques*, edited by Jonathan Harrington and Marija Tabain, 187–210. New York: Psychology Press.

Chafe, Wallace. 1982. “Integration and Involvement in Speaking, Writing, and Oral Literature.” In *Spoken and Written Language: Exploring Orality and Literacy*, edited by Deborah Tannen, 35–53. Norwood, NJ: Ablex.

Chomsky, Noam. 1959. “A review of B.F. Skinner’s *Verbal Behavior*.” *Language* 35 (1): 26–58.

Collot, Milena and Nancy Belmore. 1996. “Electronic language: A new variety of English”. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 13–28. Amsterdam: Benjamins.

Condon, Sherri and Claude Čech. 2010. “Discourse Management in Three Modalities.” *Language@Internet* 7, article 6.

http://www.languageatinternet.org/articles/2010/2770, accessed January 24, 2013.

Coulmas, Florian. 1989. *The Writing Systems of the World.* Oxford: Blackwell.

Croft, William. 2000. *Explaining language change: an evolutionary approach.* Harlow: Longman.

Dahlberg, Lincoln. 2004. “Internet Research Tracings: Towards Non-Reductionist Methodology.” *Journal of Computer-Mediated Communication* 9(3) http://jcmc.indiana.edu/vol9/issue3/dahlberg.html

Danet, Brenda and Susan Herring. 2007. *The Multilingual Internet*. Oxford: OUP.

Deutscher, Guy. 2000. *Syntactic change in Akkadian: the evolution of sentential complementation*. Oxford: OUP.

Fought, John, Robert Munroe, Carmen Fought, and Erin Good. 2004. “Sonority and Climate in a World Sample of Languages: Findings and Prospects.” *Cross-Cultural Research* 38 (1): 27–51.

Fung, Loretta and Ronald Carter. 2007. “Cantonese e-discourse: A new hybrid variety of English.” *Multilingua* 26: 35–66.

Gao, Liwei. 2006. “Language contact and convergence in computer-mediated communication.” *World Englishes* 25 (2): 299-308.

Garcia, Angela, and Jennifer Jacobs. 1999. “The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication.” *Research on Language & Social Interaction* 32 (4): 337–367.

Gee, James. 1986. “Orality and Literacy: From *The Savage Mind* to *Ways With Words.*” *TESOL Quarterly* 20 (4): 719–746.

Georgakopolou, Alexandra. 2006. “Postscript: Computer-mediated communication in sociolinguistics”. *Journal of Sociolinguistics* 10 (4): 548–557

Gibson, James. 1977. “The theory of affordances.” In *Perceiving, acting, and knowing*, edited by R. E. Shaw & J. Brandsford, 67–82. Hillsdale, NJ: Lawrence Erlbaum.

González-Lloret, Marta. 2011. “Conversation Analysis of Computer-Mediated Communication.” *CALICO Journal* 28 (2): 308–325.

Greiffenstern, Sandra. 2009. The Influence of Computers, the Internet and Computer-Mediated Communication on Everyday English. Berlin: Logos Verlag.

Haspelmath, Martin. 2010. “Framework-free grammatical theory.” In *The Oxford handbook of grammatical analysis*, edited by Bernd Heine and Heiko Narrog, 341–365. Oxford: OUP.

Hård af Segerstad, Ylva. 2002. “Use and Adaptation of Written Language to the Conditions of Computer-Mediated Communication.” PhD diss., University of Gothenburg.

Herring, Susan. 1999. “Interactional coherence in CMC.” *Journal of Computer-Mediated Communication* 4 (4).

Hutchby, Ian. 2001. *Conversation and Technology: From the Telephone to the Internet*. Cambridge: Polity.

Itkin, Il′a. 2007. *Russkaâ morfonologiâ* [Russian morphonology] Moscow: Gnozis.

Jarbou, Samir and Buthaina al-Share. 2012. “The Effect of Dialect and Gender on the Representation of Consonants in Jordanian Chat.” *Language@Internet* 9, article 1. http://www.languageatinternet.org/articles/2012/Jarbou, accessed January 24, 2013.

Jones, Graham and Schieffelin, Bambi. 2009. “Enquoting voices, accomplishing talk: Uses of *be* + *like* in instant messaging.” *Language & Communication* 29: 77–113.

Karlsson, Fred. 2009. “Origin and maintenance of clausal embedding complexity.” In *Language complexity as an evolving variable*, edited by David Gil, Geoffrey Sampson and Peter Trudgill. Oxford: OUP.

Keating, Elizabeth and Gene Mirus. 2003. “American Sign Language in virtual space: Interactions between deaf users of computer-mediated video communication and the impact of technology on language practices.” *Language in Society* 32: 693–714.

Kibrik, Aleksandr. 1992. *Očerki po obŝim i prikladnym voprosam âzykoznaniâ* [Essays on general and applied aspects of linguistics] Moscow: MGU.

Lancaster, Sean, David C. Yen, Albert H. Huang, and Shin-Yuan Hung. 2007. “The selection of instant messaging or e-mail: College students’ perspective for computer communication”. *Information Management & Computer Security* 15(1): 5–22.

LaPolla, Randy. 2003. “Why languages differ: Variation in the conventionalisation of constraints on inference.” In *Language variation: Papers on variation and change in the Sinosphere and in the Indosphere in honour of James A. Matisoff*, edited by D. Bradly, R. LaPolla, B. Michailovsky, G. Thurgood, 113–144. Canberra: Pacific Linguistics.

Lee, Carmen K.-M. 2007.” Affordances and Text-Making Practices in Online Instant Messaging.” *Written Communication* 24: 223–249.

Liberman, Mark. 2011. “Telegraphic Language” *The Language Log*, December 31. http://languagelog.ldc.upenn.edu/nll/?p=3674

Marcoccia, Michel, Hassan Atifi and Nadia Gauducheau. 2008. “Text-Centered versus Multimodal Analysis of Instant Messaging Conversation.” *Language@Internet* 5, article 7. http://www.languageatinternet.org/articles/2008/1621, accessed January 24, 2013.

Markov, Aleksandr. 2010. *Rožd*e*nie složnosti*. *Èvolûcionnaâ biologiâ segodnâ: neožidannye voprosy i novye otkrytiâ* [The emergence of complexity. Evolutionary biology today: unexpected questions and new discoveries]. Moscow: Astrel′.

Matras, Yaron. 2009. *Language Contact*. Cambridge: CUP.

Murray, Denise. 1991. *Conversation for action*: *The computer terminal as medium of communication*. Amsterdam, Philadelphia: Benjamins.

Naper, Ida. 2011. “Conversation in a multimodal 3D virtual environment.” *Language@Internet* 8, article 7. http://www.languageatinternet.org/articles/2011 /Naper, Accesses January 24, 2013.

Norman, Don. 1990. *The design of everyday things*. New York: Doubleday.

Norman, Don. 1999. Affordance, conventions and design. *Interactions* (May/June): 39-42.

Öberg, Mattias, Maritta S Jaakkola, Alistair Woodward, Armando Peruga, and Annette Prüss-Ustün. 2011. “Worldwide burden of disease from exposure to second-hand smoke: a retrospective analysis of data from 192 countries.” *The Lancet* 377 (9760): 139–146.

Olson, David. 1996. “Towards a psychology of literacy: on the relations between speech and writing.” *Cognition* 60: 83–104.

Palfreyman, David, and al Khalil, Muhamed. 2007. “‘A Funky Language for Teenzz to Use’: Representing Gulf Arabic in Instant Messaging.” In *The Multilingual Internet*, edited by Brenda Danet and Susan Herring, 43–63. Oxford: OUP.

Plester, Beverly and Clare Wood. 2009. “Exploring relationships between traditional and new media literacies: British preteen texters at school.” *Journal of Computer-Mediated Communication* 14: 1108–1129.

Rowe, Charlie. 2011. “*Whatchanade?* Rapid Language Change in a Private Email Sibling Code.” *Language@Internet* 8, article 6. http://www.languageatinternet.org/articles/2011/Rowe, accessed January 24, 2013.

Sandler, Wendy, and Lillo-Martin, Diane. 2006. *Sign language and linguistic universals.* Cambridge: CUP.

Sannikov, Vladimir. 2002. *Russkij âzyk v zerkale âzykovoj igry* [Russian language in the mirror of language play] (2nd ed.). Moscow: Âzyki slavânskoj kul′tury.

Sidorova, Marina. 2006. *Internet-lingvistika: russkij âzyk. Mežličnostnoe obŝenie.* [Internet-linguistics: Russian language. Interpersonal communication]. Moscow: 1989.ru

Skinner, B.F. 1957. *Verbal Behavior.* New York: Appleton-Century-Crofts.

Smith, Bryan and Greta Gorsuch. 2004. “Synchronous computer mediated communication captured by usability lab technologies: New interpretations.” *System* 32: 553-575.

Smyk-Bhattacharjee, Dorota. 2006. “‘A ‘lyricist’ by night, teenager by day and sleepaholic everytime in between’. Language change in/through CMC”. *Germanistische Linguistik* (186–187): 68–85.

Sparrow, Betsy, Jenny Liu, and Daniel M. Wegner. 2011. “Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips.” *Science* 333 (6043): 776–768.

Squires, Laurel. 2007. “Whats the use of apostrophes? Gender differences and linguistic variation in instant messaging.” *TESOL Working Papers* 4. Accessed

January 24, 2013 at http://aladinrc.wrlc.org/bitstream/handle/1961/5240 /CMCSquiresFinal.pdf?sequence=1

Squires, Laurel. 2010. “Enregistering internet language.” *Language in Society* 39: 457–492.

Tagliamonte, Sali, and Denis, Derek. 2008. “Linguistic ruin? LOL! Instant messaging and teen language.” *American Speech* 83 (1): 3–34.

Testelec, Âkov. 2001. *Vvedenie v obŝij sintaksis* [Introduction into the general syntax] Moscow: RGGU.

The “Five Graces Group” (Clay Beckner, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann). 2009. “Language Is a Complex Adaptive System: Position Paper”. *Language Learning* 59, s1: 1–26.

Thurlow, Crispin. 2006. “From statistical panic to moral panic: The metadiscursive construction and popular exaggeration of new media language in the print media.” *Journal of Computer-Mediated Communication* 11:3. http://jcmc.indiana.edu/vol11/issue3/thurlow.html

Thurlow, Crispin. 2007. Fabricating youth: new-media discourse and the technologization of young people. *Language in the Media*, edited by S. Johnson and A. Ensslin, 213–233. London: Continuum.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: OUP

Vandergriff, Ilona, and Carolin Fuchs. 2009. “Does CMC Promote Language Play? Exploring Humor in Two Modalities.” *CALICO Journal*, 27(1). https://calico.org/a-778-Does%20CMC%20Promote%20Language%20Play%20Exploring%20Humor%20in%20Two%20Modalities.html

Varnhagen, Connie, G. Peggy McFall, Nicole Pugh, Lisa Routledge, Heather Sumida-MacDonald, and Trudy E. Kwong. 2010. "lol: new language and spelling in instant messaging" *Reading and Writing* 23 (6): 719–733.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. *Empirical foundations for a theory of language change.* Austin: University of Texas Press.

Werry, Christopher. 1996. “Linguistic and interactional features of Internet Relay Chat. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 47–61. Amsterdam: Benjamins.

Wray, Alison and George Grace. 2007. “The consequences of talking to strangers: evolutionary corollaries of socio-linguistic influences on linguistic form.” *Lingua* 117 (3): 543–578.

Yang, Chunsheng. 2007. “Chinese Internet language: A sociolinguistic analysis of adaptations of the Chinese writing system.” *Language@Internet* 4, article 2. http://www.languageatinternet.org/articles/2007/1142. Accessed January 24, 2013.

Yates, Simeon. 1996. “Oral and Written Linguistic Aspects of Computer Conferencing: A Corpus Based Study” In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 29–46. Amsterdam: Benjamins.

Zaliznâk, Andrej. 2004. *Drevnenovgorodskij dialekt*. [Old Novgorod dialect] Moscow: Âzyki slavânskoj kul′tury.

Zitzen, Michaela, and Dieter Stein. 2004. “Chat and conversation: a case of transmedial stability?” *Linguistics* 42 (5): 983–1021.

II

2. E-mail vs. chat: the influence of the communication channel on the language

1. Introduction

Linguists are paying ever increasing attention to computer-mediated communication (CMC) and "the language of the Internet". At the "Dialogue" conference[6] the Internet is usually viewed as a tool, not an object of linguistic research; however, even here one can find papers that focus on the linguistic properties of electronic communication (Makarov and Školovaâ 2006; Zaliznâk and Mikaèlân 2006; Buras and Krongauz 2007; Bogdanov 2008; Anni 2008; Zanegina 2009; Lûdovik 2010). In order to pursue a study of this kind, the scholar has to assume that the linguistic properties of CMC are somewhat different from those of other media (oral speech and written speech, for example) and thus worthy of separate research.

This assumption is often based on a more general one: the physical properties of the communication channel affect the linguistic properties of communication taking place in this channel, acting either as constraints or as affordances. This hypothesis has been well researched in the context of the differences between written and oral speech:e.g. see the classic works of Chafe (1982) and Biber (1988). Later, interest in this field was reinvigorated by the emergence and spread of a new channel, namely CMC. The constraints there seem to be heavier than in "traditional" channels, and the affordances wider, so that one might expect that their influence on the language would be clearly visible and detectable by quantitative methods.

Since the 1980s there have been quite a few studies that have used quantitative approaches to examine differences and similarities between CMC and other channels. It is important to keep in mind that CMC is not monolithic, and that we are in fact speaking about a set of different communication channels, united by the same

[6] A major conference on computational and general linguistics held annually in Russia.

physical medium: these channels have been compared to each other as well. See Collot and Belmore 1996, Yates 1996, Hård af Segerstad 2002 and review therein, Jensen 2008, Ling and Baron 2007, Tagliamonte and Denis 2008 and review therein. The results show that CMC (or rather the specific channel studied — instant messaging, e-mail, computer conferencing and so on) is indeed a new linguistic register, neither oral speech nor written speech, and often looks like a hybrid of these two. Asynchronous communication channels with unlimited buffer size (e.g. e-mail) tend to be more similar to traditional written speech, whereas quasi-synchronous channels (see Chapter 5 about the notion of *quasi*-synchronicity), especially with limited buffer size (instant messaging), are more similar to oral speech. However, Ko (1996) shows that, in certain parameters, CMC is even more "spoken" than speech and more "written" than writing.

2. Aim of this study

My intention is to compare two communication channels within CMC: e-mail and a certain type of instant messaging. A principal novelty of this study is that the registers compared differ just by one parameter, namely the communication channel, whereas all other parameters (communicators and their relation to each other, subject matter, time of the discussion etc.) are controlled for as much as possible.

The studies mentioned above are often criticized precisely because of the lack of a control for additional parameters. Critics claim that the differences ascribed to the influence of the communication channel might in reality depend on other factors, e.g. the subject of discussion. Androutsopoulos (2006) takes this criticism even further: he states that the focus of attention should be the social context of a discourse and not its channel-specific properties. He even raises doubts about the existence of any linguistic features which might be ascribed to a communication channel: "It is empirically questionable whether in fact anything like a 'language of e-mails' exists, simply because the vast diversity of settings and purposes of e-mail use outweigh any

common linguistic features" (Androutsopoulos 2006, 420). It is not obvious, however, that the answer to this empirical question is "no". Here, I intend to address a narrower question: *does the communication channel per se have any influence on the linguistic properties of communication?*

Another novelty of my study is that I am analyzing Russian: it seems important to take CMC studies beyond the Anglophone world.

3. Materials

As a data source, I am using the contents of my own Gmail mailbox. Gmail provides not only the usual e-mail communication, but also a chat system (called Gmail chat). Since the chat is integrated into the same window (Fig. 1) and is easy to use, it is becoming increasingly popular.

Hence, it is common for the same two people to communicate both via e-mail and via chat. I collect my chat and e-mail conversations with three persons from my contact list. In order to avoid the observer's paradox, I am only using conversations which took place after June 2007 (after I graduated) and before March 2009 (before I submitted a proposal for my current PhD position), that is, when I was neither studying nor working as a linguist and did not have an idea of the current study (or anything similar) in mind.

This selection of material allows me to control for all the parameters except the communication channel itself. Indeed, the interlocutors are always the same, the setting is always the same, the subject matter may, of course, vary, but in general the same things are discussed in both chat and in e-mail messages. There is no distribution of topics (such as chat for personal matters, e-mail for business). Conversation topics include mostly personal, business, scholarly and educational matters, and none of these classes is restricted to a particular channel.

There are four subjects in my corpora: all male, native speakers of Russian, at the moment of communication aged 18 to 32, one university student, two journalists and one researcher. The e-mail corpus consists of 12 260 words and the chat corpus of 17 671 words, giving 29 931 words in total. The communication is always one-to-one.

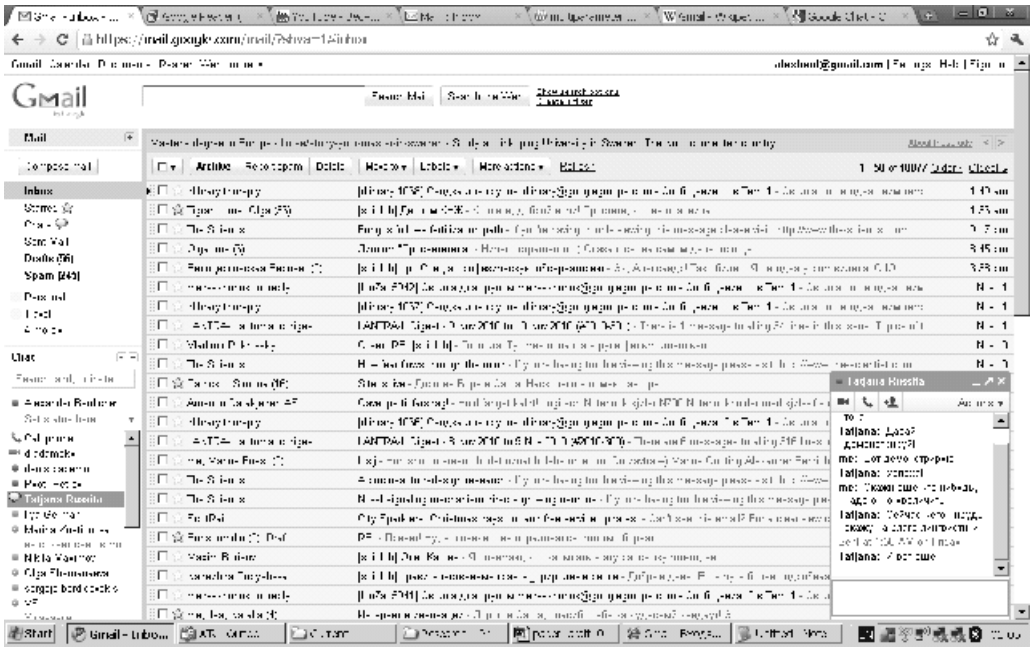

Figure 1. Gmail chat. The Chat window is in the bottom right-hand corner. The contacts list can be seen in the bottom left-hand corner, and the name of the person who has sent a new message is highlighted.

4. Methods

Biber (1994) outlines a framework for the comparison of two registers. The framework consists of three components: analysis of the situational characteristics of the registers, analysis of the linguistic characteristics of the registers, and analysis of the functional and conventional associations between situational and linguistic characteristics. This section includes the situational analysis and lists the parameters for linguistic analysis . The “Results” section provides the results of the comparison of these parameters . The “Conclusions” section discusses the associations between

situational and linguistic characteristics . This approach might be viewed as behavioral reductionism: I try to look at the influence of simple situational parameters on linguistic behavior.

4.1. Differences between the situational characteristics of e-mail and Gmail chat

First, chat messages are delivered instantly. E-mails are also delivered quickly, but it might take a few seconds (or even minutes) for a letter to reach its addressee.

Second, when you type an e-mail, your text is being auto-saved on a regular basis, so you do not have to worry about losing it should your browser crash, your Internet connection be lost, or your computer stop working. When you type a message in a chat window, it is not saved anywhere until you send it.

Third, the chat window is narrow and small (see Fig. 1), while e-mail can occupy almost the whole screen. It is possible to open the chat in a separate window (and make it as large as one wants), or to install additional software in order to make chatting more convenient, but my subjects typically use the basic small window.

Fourth, when your interlocutor is typing a chat message to you, you can see an info message “XXX is typing...” (or “XXX has entered text”) in the chat window.

Fifth, chat is more prone to technical failures: messages are more likely to get lost.

These are the real and primary differences between the two channels. They lead to the emergence of numerous secondary differences. For instance, in theory you may use chats to write long complex texts, but in practice it is awkward: first, you always risk losing everything you have typed, second, it is inconvenient to read (and type, and edit) a large text in a small window. Some of these secondary differences are not, in fact, driven by physical reality, they are conventional. Strictly speaking, you do not have to reply to chat messages immediately, but that is what you are expected to do

and what you usually do (and info messages contribute to users staying online and waiting for a reply to come).

Thus, chatting is usually a more synchronous, faster form of communication, implying immediate responses and rapid changes of speaker. It is also somewhat less reliable and more volatile.

According to Biber, one of the principal oppositions in register comparison is informational versus involved production: "discourse with interactional, affective, involved purposes, associated with strict real-time production and comprehension constraints, versus discourse with highly informational purposes, which is carefully crafted and highly edited" (1988: 115). Oral speech is usually located closer to the "involved" pole of this dimension, while written speech — closer to the "informational" pole. It seems natural to expect that the same would be true forchat and e-mail respectively. Thus, many of the linguistic parameters discussed below are those that allow one to estimate the position of a register on this scale.

4.2. Quantitative parameters for discovering linguistic characteristics of e-mail and chat[7]

1. Mean length of an utterance (MLU)

Utterance here means 'sentence', with one exception: in chat, each message is considered a separate utterance, i.e. a message might consist of several utterances (=sentences), but not vice versa. If a user chooses to split one sentence into nine messages (this is known to happen, although in my corpus such examples are rare), they are counted as nine utterances.

[7] Qualitative differences are not analyzed in this study.

Otherwise, periods, exclamation, interrogation and ellipsis marks as well as emoticons were considered as marks to end an utterance. MLU is measured in symbols. High MLU is typical of informational speech production.

2. Mean length of a word (MLW)

High MLW is typical of informational speech production. Ko (1996) finds MLW to be equal in speech and in instant messaging, but different from that in writing.

3. Lexical density (LD)

The ratio of lexical items (nouns, adjectives, verbs, adverbs, pronouns, numerals, as opposed to conjunctions, interjections, particles and prepositions) to the total number of words in a text. High LD is typical of informational speech production. Yates (1996, 35–39) shows that the LD of computer conferencing is close to that of writing, although still significantly different.

4. Type/token ratio (TTR)

The ratio of *different* words (types) in the text to the total number of words (tokens) in a text. Different word forms of the same lexeme were considered the same type, but different tokens. This measure depends on the text length, so it was calculated using two sub-corpora of equal size: 4 000 words.

High TTR implies a rich vocabulary and is typical of informational speech production. Yates (1996, 33–35) shows that the TTR of computer conferencing is close to that of writing, although still significantly different.

5. Sentence end marks

The percentage of sentences with any visible end marks: period, exclamation, interrogation or ellipsis marks. Sentences ending with an emoticon were also considered to have an end mark: sentence end is the most typical position for emoticons, and the period is usually omitted before them, so they can be viewed as an explicit signal of sentence end.

6. Capitals

The percentage of sentences beginning with a capital letter, as required by the rules of Russian punctuation/orthography.

7. Personal pronouns (first person, singular)

The ratio of the number of occurrences of the pronoun *я* ('I, me') (in all its forms) to the total number of words in a text. A high ratio is typical of "involved" speech production. Yates (1996, 40–41) finds that the proportion of first-person pronouns in total pronoun use in CMC is higher than in speech, and in speech higher than in writing. Tagliamonte and Denis (2008, 16) confirm the first part of this finding for instant messaging.

8. Brackets

The ratio of the number of brackets to the total number of words in the text. Brackets, too, serve as an indicator of informational production: a complex embedded structure (both semantic and syntactic) is difficult to create (and perceive) when text is produced (and read) "on the fly".

9. Emoticons

The ratio of the number of emoticons to the total number of words in the text. The functional spectrum of emoticons is quite broad, but it seems safe to state that speakers often use them to compensate for the lack of non-verbal cues. Thus, high emoticon ratio would imply higher involvement.

10. Complex sentences

The ratio of complex sentences (i.e. sentences containing more than one clause) to the total number of sentences. Complex sentences are typical of informational production. This measure could not be calculated automatically, so it was calculated manually using the same sub-corpora that were compiled for measuring TTR.

5. Results

The results are summarized in Table 1. All the parameters were computed for each person and each pair separately, but only the results for the whole corpus are reported, since patterns were nearly the same in all cases. The results of significance testing are reported, as well as effect sizes[8]..

a)	**Utterance length**	**Sentence end marks (%)**	**Capitals (%)**	**Complex sentences (%)**	**Emoticons (‰, 1 per 1000)**
Chat	33.8	54.7	78.3	22.9	22.9
E-mail	56.5	98.0	97.3	42.9	7.3
Δ	22.7	43.3	19,0	20.0	15.6
Significan	yes*	yes	yes	yes	yes
Effect size	medium**	large	medium	small	small

b)	**1sg**	**LD**	**TT**	**Word**	**Brackets**
Chat	3.1	74.1	31.4	4.88	4.4
E-mail	3.0	75.1	30.2	5.03	9.3
Δ	0.1	0.9	1.2	0.15	4.9
Significan	no	no	no	yes	yes
Effect size	none	non	none	none	none

Table 1. a) Variables which are manifestly different for the two channels (difference is both significant and important); b) variables which are not.
*yes means p≤ 0.05 (in fact p is smaller than 0.001 in all the cases), no − p>0.05
**large means h>0.80, medium − h>0.50, small − h>0.10, none − h≤ 0.10
Welch two-sample t-test (two-sided) applied for MLU and MLW; two-sample proportion test, for all the other cases. Effect size calculated as Cohen's d for MLU and MLW and as Cohen's h (arcsine transformation) in all the other cases.

[8] Significance testing shows how likely it is that the observed effect is random. It does not show how large and important it is. Since large samples can make very small effects visible, it is becoming increasingly common to report not only traditional significance, but also effect size (APA 2010: 33, Perry 2005: 224).

6. Conclusions

Since five parameters appeared to be truly different for e-mail and chat, we can give a positive answer to the main research question: *yes, the communication channel can influence the language*.

The sentences are shorter in chat, due to the higher speed of communication: since an immediate answer is expected, people try to be quick rather than elaborate, and do not waste much time on editing and improving their texts (especially given that chat is not the best place to do that). Interestingly, that does not affect word lengths: the pressure is probably not strong enough to make that happen.

The lack of sentence end marks and capital letters occurs for two reasons. First of all, the need for speed leads to a weakening of the norm. Second, the norm actually turns out to be unnecessary: if a message contains only one sentence (and that is usually the case), then even without capitals and periods it is clear where the sentence begins and where it ends. It would be different in a letter or in a message containing several sentences, but in these cases the norm is usually not ignored.

It might also be supposed that chat is considered to be a less formal channel where norm violations are more appropriate, but this claim is hard to prove or disprove using my data.

Emoticons are more numerous in chat, since in a quasi-synchronous mode it is more important to show a "polite smile" to an interlocutor. They also have a phatic function: you are showing that you are interested in what your partner is saying, and you might reply to a message with a single smiling emoticon if you do not have anything else to say.[9]

[9] As one of the subjects of this study put it, when questioned, "...I also want to be polite, so in chat I actually use a smiley instead of a period :)".

It is interesting to compare my results to those reported in Baron 2004 for English instant messaging (IM). Baron finds 49 instances of emoticons in her corpus of 11 718 words (Baron 2004, 413), the ratio being 0.004. My ratios are 0.023 (405/17671) for chat, 0.007 (90/12260) for e-mail, and 0.017 counted together — that is, much higher. This is unexpected, since the participants in my study are older and more educated than in Baron's. Besides, Baron's sample includes female subjects, and women tend to use more emoticons than men (Baron 2004, 416). It is unlikely that Russian IM is so much richer in emoticons than English IM. One explanation might be the observer's paradox: Baron's subjects knew they were being recorded while chatting, and this might easily have influenced their speech production (they might have tried to avoid "informal" traits like emoticons). Alternatively, it is possible that emoticons were less popular when Baron's study was conducted.[10]

For brackets, the difference is significant, but the effect size is too small. Most likely this means that there actually is a difference, but the sample is too small to show it.

As for the other parameters, we might be quite sure that there are no differences, or that they are really tiny. This means that the influence of the communication channel should not be overestimated.

In other words, my data show that the change of channel can trigger differences in choice of syntactic structures, discourse and pragmatic devices, as well as in the level of adherence to norms. The observed differences are generally in line with previous findings on e-mail and chat, and their emergence can be explained by means of primary and secondary affordances of the channels. This does not mean, however, that these differences will always *necessarily* arise: as Androutsopoulos (2006) correctly observes, there exist many other factors which can also influence linguistic structure. The conclusions should be rather viewed as probabilistic: in chat, it is more *probable* that the sentences will be shorter on average than in e-mail, etc.

[10] This possibility was suggested to me by Aleksandr Piperski.

Further development of this study might include analysis of more complex parameters and data from the other social groups: less educated, less language-aware and not including myself, the researcher. It would also be useful to compare another set of channels, but it would be difficult, if not impossible, to reduce the distinction between registers to this single parameter.

References

American Psychological Association. 2010. *Publication manual of the American Psychological Association, 6th edition*. Washington, DC: APA.

Androutsopoulos, Jannis. 2006. "Introduction: Sociolinguistics and computer-mediated communication." *Journal of Sociolinguistics* 10 (4): 419–438.

Anni, Oja. 2008. "Choosing language in Internet conversations between Russians and Estonians." In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"* 7 (14), 602–605. Moscow: RGGU.

Baron, Naomi. 2004. "See You Online: Gender Issues in College Student Use of Instant Messaging." *Journal of Language and Social Psychology* 23: 397–423.

Biber Douglas. 1998. *Variation across speech and writing*. Cambridge: CUP.

Biber, Douglas. 1994. "An Analytical Framework for Register Studies". In *Sociolinguistic perspectives on Register*, edited by Douglas Biber and Edward Finegan, 44–56. New York: OUP.

Bogdanov, Aleksej. 2008. "Orfografiâ v Internete: analiz odnoj orfografičeskoj ošibki." [Ortography at the Internet: the analysis of one orthographical error] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"* 7 (14), 50–56. Moscow: RGGU.

Buras, Mariâ, and Maksim Krongauz. 2007. “Âzyk korporativnyh sajtov: igra, parodiâ, provokaciâ.” [The language of the corporate websites: play, parody, provocation] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference “Dialogue”*, 109–114. Moscow: RGGU.

Chafe, Wallace. 1982. “Integration and Involvement in Speaking, Writing, and Oral Literature.” In *Spoken and Written Language: Exploring Orality and Literacy*, edited by Deborah Tannen, 35–53. Norwood, NJ: Ablex.

Collot, Milena and Nancy Belmore. 1996. “Electronic language: A new variety of English”. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 13–28. Amsterdam: Benjamins.

Hård af Segerstad, Ylva. 2002. “Use and Adaptation of Written Language to the Conditions of Computer-Mediated Communication.” PhD diss., University of Gothenburg.

Jensen, Bård Uri. 2008. “Syntactic variables in pupils’ writing: A comparison between hand-written and pc-written texts.” In *Linguistics in the making: selected papers from the second Scandinavian PhD conference in linguistics and philology in Bergen, June 4-6, 2007*, edited by Gard B. Jenset, Øystein Heggelund, Margrete Dyvik Cardona, Stephanie Wold, and Anders Didriksen, 165–184. Oslo: Novus Press.

Ko, Kwang-Kyu. 1996. “Structural Characteristics of Computer-Mediated Language: A Comparative Analysis of InterChange Discourse.” *Electronic Journal of Communication* 6 (3)

Ling, Richard, and Naomi Baron. 2007. “Text Messaging and IM: Linguistic Comparison of American College Data.” *Journal of Language and Social Psychology* 26 (3): 291–299.

Lûdovik, T. 2010. Analiz tekstov SMS-soobŝenij s celʹû povyšeniâ kačestva [Analysis of SMS texts aimed at improving their automatical reading] // In

Computational linguistics and intellectual technologies. *Papers from the annual international conference "Dialogue"* 9 *(*16), 313–317. Moscow: RGGU.

Makarov, M., and M. Školovaâ. 2006. "Lingvističeskie i semiotičeskie aspekty konstruirovaniâ identičnosti kommunikacii." [Linguistic and semiotic aspects of identity construction in e-communication] *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"*, 364–369. Moscow: RGGU.

Perry, Fred. 2005. *Research in applied linguistics: becoming a discerning consumer*. Mahwah: Laurence Erlbaum Associates.

Tagliamonte, Sali, and Derek Denis. 2008. "Linguistic ruin? LOL! Instant messaging and teen language." *American Speech* 83 (1): 3–34.

Yates, Simeon. 1996. "Oral and Written Linguistic Aspects of Computer Conferencing: A Corpus Based Study" In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 29–46. Amsterdam: Benjamins.

Zaliznâk, Anna, and Irina Mikaèlân. 2006. "Perepiska po èlektronnoj počte kak lingvističeskij ob″ekt." [E-mail correspondence as an object of linguistic analysis] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"*, 157–162. Moscow: RGGU.

Zanegina, N. 2009. "~~Â ètogo ne govoril~~: o liturarivah, začerkivaniâh ili mnimyh tekstah." [~~I've never told that~~: about lituratives, strikeouts or imaginary texts] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"* 8 (15), 112–115. Moscow: RGGU.

III

3. The 'orthographic' neuter: a grammatical innovation in Russian Internet language

1. Introduction

In recent times linguists have spoken more and more about the phenomenon of 'Internet language'. The paper that preceded the writing of this article was presented at a round table called "Internet communication, mass media, new means of communication". To consider Internet language a subject worthy of special research,[11] it is necessary to postulate that it is in some ways different to other speech registers.[12]

Quantitative differences between registers have been convincingly demonstrated and thoroughly studied (for examples see Tagliamonte and Denis 2008, with bibliography). For the Russian language similar issues are analyzed in Chapter 2, with special emphasis on the function of the communication channel.

It would be even more remarkable to find examples of *qualitative* influence, i.e. systemic innovations in language that are induced by the new channel. The present article describes one such innovation: a mass-scale migration of nouns into neuter gender under the influence of anti-standard orthography.

Readers not familiar with the Russian language are invited to review its relevant properties in Appendix 1.

[11] Krongauz (2008) is correct when he says that "there is no Internet language as such, or, rather, it is just as diverse as our everyday language". In certain cases, however, the generalization "Internet language" might be used.

[12] I use the term 'register' as described by Biber (1994): the variety of the language used in a certain situation. The situation is defined by a number of parameters (addresser, addressee, place, time, physical communication channel, etc.). The register can be defined with more precision or less; thus, one can speak of the register

2. Description of the phenomenon

2.1. Beginnings of the innovation

In the early 2000s, Runet (the Russian-language section of the Internet) was quickly swamped with an anti-standard idiom called *âzyk padonkaff* (lit. 'the language of scumbags', also known as 'Olbanian language'). This phenomenon has been reviewed from various standpoints (see the linguistic description in Hristova 2011, or a brief history and meta-linguistic analysis in Zvereva 2009, as well as diachronic description in Chapter 6). For the purposes of this article, however, just one of its orthographic manifestations is relevant: the spelling of unstressed *a*. The principle 'break all rules of orthography' demands that *o* should be written instead of the unstressed *a*. Thus, nonneuter declension[13] words with nstressed ending - *дéвушка, кни́жка, нау́ка, мужчи́на* / *dévuška, knížka, nauka, mužčína* / 'girl', 'book', 'science', 'man' - assume the form *дéвушко, кни́жко, нау́ко, мужчи́но*[14] / *dévuško, knížko, nauko, mužčíno* / 'girl-oh', 'book-oh', 'science-oh', 'man-oh'. [15]

Words ending in -*o* in Russian are strongly linked to the neuter gender. In the Grammar dictionary (Zaliznâk 1977), there are 1,596 nouns ending in -*o*, 1,569 of them (98%) neuter.[16] As a consequence of this link, words like *дéвушко*/ *dévuško* / 'girl-oh' are reanalyzed as neuter, with their declension and agreement following suit.

"Internet language" or the registers "email language", "chat room language", etc. This article mostly deals with the registers of the Internet language that are used for informal communication: blogs, forums, chat rooms, etc.

[13] For the definitions of declensions and choice of labels see Appendix 1.

[14] Obviously, various orthographical deviations may occur in other positions as well: *дéвужко / dévužko, ноу́ко / поу́ко*, but they are of no interest for the purposes of this article.

[15] Realizing that functional translation of a linguistic concept into examples in a language different to that of the phenomena described is impossible, we adopted the convention of adding the '-oh' ending to simulate the subject of this article. While it fails to convey the grammatical intricacies of the Russian language's case and accentological system, it retains some of the playfulness inherent in the use of this device in Russian Internet slang.

[16] The other 27 nouns belong to the masculine gender, 25 of them with the diminutive suffix -*ишко / -iško*.

We are witnessing a rare phenomenon: orthography influencing morphology. Nouns are migrating in droves to the new 'orthographic' neuter gender (it is important to note that it does not happen every time, and the spelling of the *дéвушко*/ *dévuško* / 'girl-oh' type does not necessarily mean that the word has changed its gender). The mass nature of this migration is shown in section 4.

2.2. Declension

One could isolate three basic types of declension among new *-o*-formations. They can be declined as regular nouns of the neuter gender:

(1)[17]

Илья, вот объясни мне, тупому блонди́нку[18]*: ЧТО ТАМ СЛОЖНОГО?!* (forum, 2007)

Il'̲â, vot ob"âsni mne, tupomu blondínku: ČTO TAM SLOŽNOGO?!

'Ilya, could you explain to me, dumb-oh blonde-oh that I am, WHAT'S SO DIFFICULT ABOUT THAT?'

EXPECTED STANDARD FORMS for 'to the dumb blonde':

туп-о́й *блонди́нк-е*

[17] Examples are quoted without any changes to grammar or punctuation unless stated otherwise. The transliteration system used is ISO 9 (GOST 7.79—2000). Key words are underlined, omissions are marked by suspension points. The type of the source is given in all cases, the year is given whenever possible; if necessary, the meaning is explained in the text or in a footnote. Key forms are usually glossed (both the used variant and the expected standard variant are given) or commented on separately. Detailed glossing is provided only for morphemes which are directly relevant to the "new neuter" phenomena, i.e. mostly endings and some suffixes. Glosses are in accordance with the Leipzig Glossing Rules (http://www.eva.mpg.de/lingua/resources/glossing-rules.php). For nouns that migrated into the "new neuter", stress is marked. The examples in this article abound with deviations from standard usage, but only those that are relevant to the subject under discussion are explained and translated.

[18] Strictly speaking, it is impossible to distinguish the neuter gender from the masculine here, but the neuter is much more probable: there is one occurrence of "тупо́й блонди́нок / tupój blondínok" found, and hundreds for "тупо́е блонди́нко / tupóe blondínko".

tup-ój *blondínk-e*
'dumb-DAT.SG.F blonde-DAT.SG.D1'

USED FORMS:

туп-о́му *блонди́нк-у*
tup-ómu *blondínk-u*
'dumb-DAT.SG.N blonde-DAT.SG.D2'

They can lose declension altogether, both in singular form (example 2) and in plural (example 3). The main reason for indeclinability is probably that such words are construed as analogous to numerous indeclinable nouns ending in -o, but other factors may be at play as well (see below).

(2)
Пылин, завтра у моей дéвушко[19] *бéзднико.*

Pylin, zavtra u moej dévuško bézdniko.

'Damn, my girlfriend-oh has her birthday-oh tomorrow' (blog, 2005)

EXPECTED STANDARD FORMS FOR 'my girlfriend has':

y *мо-ей* *дéвушк-и*
u *mo-ej* *dévušk-i*
'at my-GEN.SG.F girlfriend-GEN.SG.D1'

USED FORMS:

y *мо-ей* *дéвушко*
u *mo-ej* *dévuško*
'at my-GEN.SG.F girlfriend.D0'

[19] Many examples contain the word *dévuško* 'girl'. The only reason is that this was one of the first "new neuter" forms I encountered, and I used it as a keyword to search for examples.

(3)
Андрюх, мне кажется ба́бушко нашего возраста, ну может чуть чуть постарше... но мы то не де́душко! (forum, 2007)

Andrûh, mne kažetsâ bábuško našego vozrasta, nu možet čut' čut' postarše... no my to ne déduško!

'Andrûh, I think the granny-oh is about as old as us, maybe a tiny bit older… but us, we're not grandfathers-oh!'

EXPECTED STANDARD FORMS FOR 'but us, we are not grandfathers':

мы=то	*не*	*де́душк-и*
my=to	ne	*dédušk-i*
'my=but	not	grandfather-NOM.PL.D1'

USED FORMS:

мы=то	*не*	*де́душко*
my=to	ne	*déduško*
'my=but	not	grandfather.D0'

Finally, the original declension might be preserved in full:

(4)
Де́вушко уже взрослое - будет 4 года. (...) Понос! мама (я) и бабушка обучают де́вушку правилам безопасности и гигиены во время поноса. (...) Де́вушка чувствует себя отлично (...) Де́вушке разрешено на папе играть и отрабатывать гаммы. Де́вушко танцует на заднице полупроснувшегося папы (...) (bash.org.ru, 2008)

Dévuško uže vzrosloe - budet 4 goda. (...) Ponos! mama (â) i babuška obučaût dévušku pravilam bezopasnosti i gigieny vo vremâ ponosa. (...) Dévuška čuvstvuet sebâ otlično (...) Dévuške razrešeno na pape igrat' i otrabatyvat' gammy. Dévuško tancuet na zadnice poluprosnuvšegosâ papy (...)

'The girl-oh is already grown-up: she'll be four soon. (…) Diarrhea! Mum (me) and grandma are teaching the girl the rules of safety and hygiene during diarrhea. (…) The girl feels great. (…) The girl is allowed to play and rehearse gamuts on her dad. The girl-oh dances on the ass of her half-awake dad (…)' (bash.org.ru, 2008)

COMMENTARY: In three out of five occurrences of the word *dévuško*, the standard forms are used: *dévuška* (nom.sg), *dévušku* (acc.sg), *dévuške* (dat.sg). In two cases, however, *devuško* is used instead of nom.sg.

In example 4, the author inconsistently writes either *дéвушко / dévuško /* 'girl-oh' (and even *взрослое дéвушко / vzrosloe dévuško /* 'grown-up-oh girl-oh') or *дéвушка / dévuška /* 'girl'. Such variability is typical for virtually all manifestations of the innovation in question. Apparently speakers do not have in mind any strict rules to the effect of 'I always decline words in accordance with their original gender' or 'I always decline words in accordance with their new gender'.

The unsystematic nature of the innovation is non-random: oblique cases present a certain problem for speakers wishing to use the oh-word.[20] For nonfeminine declension nouns they are in the position of neutralization (compare *моё сту́ло /моё stúlo /* 'my-oh chair-oh' and *моему сту́лу / moemu stúlu /* 'to my chair' OR 'to my-oh chair-oh'). For nonneuter declension nouns, indirect cases may create an impression of masculine gender when the idea was to use neuter (*красивому мужчи́ну / krasivomu mužčínu*[21] */* 'to the handsome man' OR 'to the handsome-oh man-oh'). In both cases, a possible solution could be to drop declension altogether, as in examples 2 and 3. In the case of a-declension, to avoid the confusion with the masculine, the speaker may revert to the original feminine gender. This might be happening in example 4, as well as in examples 7 and 10, where the anaphoric pronoun, which has an *oh*-word as the antecedent, is used in neuter gender if it is nominative but in feminine if it is in an oblique case.

[20] I owe this observation and the subsequent analysis to Lûdmila Fëdorova.

[21] The correct form would be *красивому мужчине / krasivomu mužčine*.

2.3. Agreement

Agreement in neuter gender, observed in examples 1 and 4, makes a much stronger case for the migration to neuter gender than declension; it is also more interesting linguistically.

This parameter also shows a high degree of variability.

(5)
Я вот не в пример более видный мужчи́но, а при случае и в репу могу дать огрызку какому, чтоб не напяливал фура́жко на своё пустое ты́кво. (comment on photo hosting site, 2008)

Â vot ne v primer bolee vidnyj mužčíno, a pri slučae i v repu mogu dat' ogryzku kakomu, čtob ne napâlival furážko na svoe pustoe týkvo.

'I'm a much more presentable man-oh, but should the need arise I may give a thumper to any fuckwit for pulling a service cap-oh on his-oh empty-oh dome-oh.'

EXPECTED STANDARD FORMS FOR 'well-built man':

видн-ый	*мужчи́н-а*
vidn-yj	mužčín-a
'presentable-NOM.SG.M	man-NOM.SG.D2'

USED FORMS

видн-ый	*мужчи́но*
vidn-yj	mužčíno
'presentable-NOM.SG.M	man-NOM.SG.D1'[22]

EXPECTED STANDARD FORMS FOR 'pulled on a service cap':

напяливал-ø	*фура́жк-у*
napâlival-ø	*furážk-u*
'pull_on:IPFV:PST-SG.M	service_cap-ACC.SG.D2'

USED FORMS:

напяливал-ø	*фура́жк-о*
napâlival-ø	*furážk-o*
'pull_on:IPFV:PST-SG.M	service_cap-ACC.SG.D1'

[22] When an oh-noun is used in nom.sg or acc.sg, it is impossible to say whether it belongs to D1 or to D0, since the ending is *-o* in both cases. By default, in such cases D1 is assumed.

EXPECTED STANDARD FORMS FOR ‘on his(-oh) empty(-oh) dome(-oh)’:

на	*сво-ю*	*пуст-ую*	*ты́кв-у*
na	*svo-ûju*	*pust-uûju*	*tý́kv-u*
‘on	POSS.REFL-ACC.SG.F	empty-ACC.SG.F	dome-ACC.SG.D2’

USED FORMS:

на	*сво-ё*	*пуст-ое*	*ты́кв-о*
na	*svo-ë*	*pust-oe*	*tý́kv-o*
‘'on	POSS.REFL-ACC.SG.N	empty-ACC.SG.N	dome-ACC.SG.D1’

(6)
есть <u>интересный</u> <u>кни́жко</u> Максима Кронгауза "Русский язык на грани нервного срыва", как раз про новомодные словечки (forum, 2010)

est’ <u>interesnyj</u> <u>knížko</u> Maksima Krongauza "Russkij âzyk na grani nervnogo sryva", kak raz pro novomodnye slovečki

‘There’s an <u>interesting book-oh</u> by Maksim Krongauz, “The Russian Language on the Verge of Nervous Breakdown”, which is specifically about hot new words.’

EXPECTED STANDARD FORMS FOR ‘interesting book’:

интересн-ая	*кни́жк-а*
interesn-aâ	*knížk-a*
‘interesting-NOM.SG.F	book-NOM.SG.D2’

USED FORMS:

интересн-ый	*кни́жк-о*
interesn-yj	*knížk-o*
‘interesting-NOM.SG.M	book-NOM.SG.D1’

These examples show three possible types of agreement of -oh-words: in accordance with the original gender (*видный мужчи́но / vidnyj mužčíno /* ‘presentable-M man-oh’[23]), in accordance with the “orthographic” neuter (*своё пустое ты́кво / svoe pustoe tý́kvo /* ‘his-oh empty-oh dome-oh’), and, as Zubova (2010, 17) calls it, “demonstrative disagreement of words across all parameters” (*интересный кни́жко*

[23] Of course, agreement in the original gender is also possible for the feminine: *маленькая же́нщино / malen’kaja žénščino /* ‘little-F woman-oh’.

/ interesnyj knížko / ‘interesting-M book-oh’). It is not quite clear to what degree such disagreement is indeed demonstrative, but it does occur much more infrequently than the first two types. For the original masculine it is almost never used (collocations of the type *моя ńапо / moja ṕapo /* ‘my-F dad-oh’*, моя мужчи́но / moja mužčína /* ‘my-F man-oh’ are extremely rare), which is unsurprising: first, the transition from feminine to masculine is always easier than the opposite direction; second, the *-o /* ‘–oh’ ending is easier to tie to the masculine gender than to the feminine.

Agreement occurs not just with the attribute, but also with the verb and the anaphoric pronoun.

(7)
На четырехполосной улице рядом с двойной сплошной дéвушко припарковалось против хода движения. Меня интересует, а как онó туда приехало? И кто продал ей права? (blog, 2009)

Na četyrehpolosnoj ulice râdom s dvojnoj splošnoj dévuško priparkovalos’ protiv hoda dviženiâ. Menâ interesuet, a kak onó tuda priehalo? I kto prodal ej prava?

‘On a four-lane street, next to the double white lines, a girl-oh parked-oh [her car] against the direction of the traffic. I wonder, how did she-oh get there? And who sold her her license?’

EXPECTED STANDARD FORMS FOR ‘a girl parked [her car]’
дéвушк-а *припарковал-а-сь*
dévušk-a *priparkoval-a-s′*
‘girl-NOM.SG.D2 park:PFV:PST-SG.F-REFL’

USED FORMS:
дéвушк-о *припарковал-о-сь*
dévušk-o *priparkoval-o-s′*
‘girl-NOM.SG.D1 park:PFV:PST-SG.N-REFL’

COMMENTARY: In this extract, there are two anaphoric pronouns referring to *dévushko* ‘girl-oh’: *onó* (nom.sg.n) and *ej* (dat.sg.f).

It has been noted many times (for examples, see Norman 2006, 154-155; Il′ina 327-328) that the formal connection between noun and attribute is stronger than between noun and verb:

(8)
Подруга подходит сейчас к холодильнику, на котором гордо восседает царской статуэткой котэ́ и, потянувшись к ней, чмокает в нос. Котэ́ неожиданно шарахается назад и сшибает с холодильника на пол задницей все контейнеры, что были на нем. Это полосатое Цу́ко явно не ожидала такой наглости со стороны холопов :D [24] (bash.org.ru, 2010)

Podruga podxodit sejčas k holodil'niku, na kotorom gordo vossedaet carskoj statuètkoj kotḗ i, potânuvšis' k nej, čmokaet v nos. Kotḗ neožidanno šarahaetsâ nazad i sšibaet s holodil'nika na pol zadnicej vse kontejnery, čto byli na nem. Èto polosatoe Cúko âvno ne ožidala takoj naglosti so storony holopov :D

'My girlfriend walks up to the fridge, upon which the cat-eh is sitting like a regal statue, and kisses her on the nose. The cat-eh suddenly recoils and knocks all the containers off the top of the fridge. The striped bitch-oh clearly wasn't expecting such insolence from the hoi polloi :D'

EXPECTED STANDARD FORMS FOR 'this striped bitch clearly was not expecting...':

Эт-а	*полосат-ая*	*су́к-а*	*явно*
Èt-a	*polosat-aâ*	*súk-a*	*âvno*
'This-NOM.SG.F	striped-NOM.SG.F	bitch-NOM.SG.D2	clearly

не	*ожидал-а*
ne	*ožidal-a*
not	expect:IPFV:PST-SG.F'

USED FORMS:

Эт-а	*полосат-ое*	*су́к-о*	*явно*
Èt-a	*polosat-oe*	*súk-o*	*âvno*
'This-NOM.SG.F	striped-NOM.SG.N	bitch-NOM.SG.D1	clearly

[24] *Котэ / kotè* / 'cat-eh' is not a nonce word, it is another widespread productive model worthy of separate study. It is quite possible that this innovation develops under the influence of the 'orthographic' neuter. *Цу́ко* / *cúko* is a widespread non-standard spelling variant of the word *сука / suka* / 'bitch'.

не *ожидал-а*
ne *ožidal-a*
not expect:IPFV:PST-SG.F'

COMMENTARY: The attribute adjective *polosatoe* is in agreement with the formal gender of the noun *Cúko* (neuter), while the verb *ožidala* is in agreement with its original semantic gender (feminine).

Another curious phenomenon is obvious from example 7: anaphoric pronouns with *дéвушко / dévuško /* 'girl-oh' for antecedent are first used in neuter gender (*онó / onó /* 'she-oh'), and then in feminine (*ей / ej /* 'to her'). The same principle is at work in example 9.

(9)
Ыыыы, щас пришло клиéнтко...
Спросило есть ли у вас жидкость для жидко-кристаллических телевизоров.
На вопрос "Зачем и пачиму?" ответила, что он потек... (bash.org.ru [chat], 2010)

Yyyy, ŝas prišlo kliéntko...
Sprosilo est' li u vas židkost' dlâ židko-kristalličeskih televizorov.
Na vopros "Začem i pačimu?" otvetila, čto on potek...

'Welp… a customer-oh just came in-oh… Asked-oh if we had liquid for LCD TV sets. When asked "Why and what for?" responded that it started leaking…'

EXPECTED STANDARD FORMS FOR 'a female customer came':
пришл-а *клиéнт-к-а*
prišl-a *kliént-k-a*
'come:PST:PFV-SG.F customer-F-NOM.SG.D2'

USED FORMS:
пришл-о *клиéнт-к-о*
prišl-o *kliént-k-o*

'come:PST:PFV-SG.N customer-N-NOM.SG.D1'[25]

COMMENTARY: verbs *sprosilo* ('asked') and *otvetila* ('responded') both agree with *klientko*, but in the former case the gender is neuter, in the latter – feminine, as is clear from the verb endings.

For the verbs *спросило / sprosilo /* 'asked' and *ответила / otvetila /* 'responded', the subject is *клие́нтко / kliéntko /* 'customer-oh', but it is dropped in both cases. The first verb is closer to the only occurrence of the subject in the text, and agrees with its formal gender (*спросило / sprosilo*); by the second occurrence the speaker has probably forgotten that he had used the -oh ending and agrees the verb with the formal gender: *ответила / otvetila.* The oh-word serves as an attractor for anaphoric pronouns and verbs with the dropped subject; further from the center it becomes more likely that they would agree in accordance with the original gender. I have not found any examples of agreement according to the formal gender after an occurrence of agreement according to the original gender. It is only possible if the oh-word is used again, as in the following example:

(10)
Онó (брюнéтко, с выражением возмущения уже готовиццо к тираде, но... Ваши крокодилы, вы их и спасайте (с). Я высказала всё, что думаю о ней, о её водительских навыках, наличии мозга (...)
Тут и подмога подоспела в лице парковщика, который тоже пылал гневом в адрес сей водятельницы (...) у брюнéтко было неповторимое выражение на лице. Я ушла в магазин, а по возвращении с удовлетворением отметила, что онó припарковало своё машинко по всем правилам. (blog, 2007)

Onó (brûnétko, s vyraženiem vozmuŝeniâ uže gotovicco k tirade, no... Vaši krokodily, vy ih i spasajte (c). Â vyskazala vse, čto dumaû o nej, o ee voditel'skih navykah, naličii mozga (...)
Tut i podmoga podospela v lice parkovŝika, kotoryj tože pylal gnevom v adres sej vodâtel'nicy (...) u brûnétko bylo nepovtorimoe vyraženie na lice.

[25] The suffix *-k-* is highly homonymous. One of the meanings is diminutive. Another (possible only if the suffix followed by a d2 ending) is deriving feminine nouns from original masculine, cf. *klientka* (f) from *klient* (m). It seems that, in the case of *kliéntko*, the suffix, followed by *-o*, has to be ascribed the function of converting masculine to neuter — something it never does in the standard language.

Â ušla v magazin, a po vozvraŝenii s udovletvoreniem otmetila, čto onó priparkovalo svoe mašinko po vsem pravilam.

'She-oh (the brunette-oh), with an expression of indignation, is preparing for a rant, but… your crocodiles, you save them (c). I said everything I thought about her, about her driving skills, the existence of her brain…
Help arrived in the form of the parking attendant, who was also enraged by this so-called driver (…) The brunette-oh's facial expression was indescribable. I went to the store, and when I returned I was satisfied to see that she-oh had parked the car quite correctly.'
COMMENTARY: the pronouns referring to *brûnétko* ('brunette-oh') appear in the following order: *onó* (nom.sg.n), *nej* (loc.sg.f), *eë* (gen.sg.f). Then the noun *brûnétko* is repeated, the neuter gender is reactivated in speaker's mind, and the next pronoun is again *onó* (nom.sg.n). Another factor in choosing between neuter and feminine here might be the case of the pronoun (nominative vs. oblique), see discussion of example 4 in section 2.2.

2.4. Diffusion of the innovation

Transition to the formal neuter gender does not only cover nouns with unstressed *-a* endings, but other classes too. Transition of nouns with the stem ending in a consonant is evident in example 2 (*бе́зднико / bézdniko /* 'birthday-oh'), as well as in the following examples:

(11)
Ребенок с сомнением посмотрел в сторону отца - немытое, заросшее admino мирно рубящееся в консоли с будущим проектом под еще более брутальный хард рок. (bash.org.ru, 2010)

Rebenok s somneniem posmotrel v storonu otca - nemytoe, zarosšee admíno mirno rubâŝeesâ v konsoli s budušim proektom pod eŝe bolee brutal'nyj hard rok.

'The child cast a suspicious glance at his father, a dirty-oh, hairy-oh admin-oh, peacefully fidgeting with a future project on the console and listening to some brutal hard rock.'

EXPECTED STANDARD FORMS FOR 'dirty, hairy system administrator, [who was] peacefully fidgeting...'

Немыт-ый	*заросш-ий*	*адми́н-ø*
Nemyt-yj	*zarosš-ij*	*admín-ø*
‘Dirty-NOM.SG.M	hairy-NOM.SG.M	admin-NOM.SG.D1

мирно	*рубящ-ий-ся*
mirno	*rubâŝ-ij-sâ*
peacefully	fidget:PTCP-NOM.SG.M-REFL’

USED FORMS:

Немыт-ое	*заросш-ее*	*адми́н-о*
Nemyt-oe	*zarosš-ee*	*admín-o*
‘Dirty-NOM.SG.N	hairy-NOM.SG.N	admin-NOM.SG.D2

мирно	*рубящ-ее-ся*
mirno	*rubâŝ-ee-sâ*
peacefully	fidget:PTCP-NOM.SG.N-REFL’

(12)
все сами строят <u>свое</u> <u>первое</u> <u>дне́</u> осени. (blog, 2005)

vse sami stroât <u>svoe</u> <u>pervoe</u> <u>dné</u> oseni.

‘Everyone constructs their-oh first-oh day-oh of the autumn’

EXPECTED STANDARD FORMS FOR ‘...construct their first day...’:

стро-ят	*сво-й*	*перв-ый*	*дéнь-ø*
stro-ât	*svo-j*	*perv-yj*	*dén'-ø*
construct-PRS.3PL	POSS.REFL-ACC.SG.M	first-ACC.SG.M	day-ACC.SG.D1

USED FORMS:

стро-ят	*сво-е*	*перв-ое*	*дн-é*
stro-ât	*svo-e*	*perv-oe*	*dn-é*
construct-PRS.3PL	POSS.REFL-ACC.SG.N	first-ACC.SG.N	day-ACC.SG.D1

(13)
надо было мне тебе смс послать, <u>тупОе</u> <u>мое</u> <u>голово́</u>! (forum, 2009)

nado bylo mne tebe sms poslat’, <u>tupOe</u> <u>moe</u> <u>golovó</u>!

‘I should have sent you an SMS, <u>my-oh</u> <u>dumb-oh</u> <u>head-oh</u>!’

EXPECTED STANDARD FORMS FOR ‘my dumb head’:

туп-ая	*мо-я*	*голов-á*
tup-aâ	*mo-â*	*golov-á*
dumb-NOM.SG.F	my-NOM.SG.F	head-NOM.SG.D2

USED FORMS:

туп-Ое	*мо-е*	*голов-о́*
tup-Oe[26]	*mo-e*	*golov-ó*
dumb-NOM.SG.N	my-NOM.SG.N	head-NOM.SG.D1

COMMENTARY: In *golová*, the ultimate [a] is stressed, so changing it to [o] is not just an orthographical phemonenon (as in previous examples with unstressed endings), but a phonological one as well (as in examples 11 in 12, where an additional [o] is added to a zero ending).

Examples of transition can be found for any formal class of nouns, even though their frequency varies (see section 4). Some formal features obviously facilitate such transition; Zubova (2010, 17) correctly points out that the suffix *–к(а) / -k(a)* is one of them. Transition is discouraged (but not rendered completely impossible, see example 12) by the stem ending with a palatalized consonant.

Morphologically, the transition can be structured in various complex ways:

(14)
...только один человек написал, что сильнее всего его оскорбляет словосочетание «Ла́мо Коматозное». (bash.org.ru [chat], 2010)

...tol'ko odin čelovek napisal, čto sil'nee vsego ego oskorblâet slovosočetanie "Lámo Komatoznoe".

'…just one person wrote that he was offended the most by the phrase "Comatose-oh Lamer-oh"'

EXPECTED STANDARD FORMS FOR 'comatose lamer':

ла́мер-ø	*коматозн-ый*
lámer-ø	*komatozn-yj*
lamer-NOM.SG.D1	comatose-NOM.SG.M

USED FORMS

Ла́м-о	*Коматозн-ое*

[26] Capitalization of O is probably a marker of emphasis.

Lám-o *Komatozn-oe*
lamer-NOM.SG.D1 comatose-NOM.SG.N

(15)
вот как объяснить маме что ее ненаглядное 15-летнее школоло́ всю ночь не могло оторваться от статей по СТО, ОТО, квант. мех., теории струн, а не смотрело хентай? (bash.org.ru [chat], 2010)

vot kak ob'ʻâsnit' mame čto ee nenaglâdnoe 15-letnee škololó vsû noč' ne moglo otorvat'sâ ot statej po STO, OTO, kvant. meh., teorii strun, a ne smotrelo hentaj?

'how can one explain to mum that her dearest-oh 15-year-old-oh schoolchild-oh could not tear itself away from articles on general and special relativity, quantum mechanics, string theory, and did not watch hentai?'

Ла́мо / lámo / 'lamer-oh' is formed from *ламер* / 'lamer' ('computer illiterate') using an apocope untypical for Russian. *Школоло́ / škololó* is formed by the contamination of the pejorative collective noun *школота́ / školotá* / 'schoolchildren' and a popular interjection *ололо́ / ololó*.[27]

2.5. Speakers' attitude towards the innovation

The linguistic function of this innovation cannot be fully described without evaluating how its use is perceived by speakers. Explicit metacomments are infrequent:

(16)
с де́вушко слушай, ещё раз напишешь меня в среднем роде и выход тебе точно уже не понадобится) (blog, 2010)

[27] Possibly influenced by Eduard Hil'ʼs song, "Trololo".

s dévuško slušaj, eŝe raz napišeš' menâ v srednem rode i vyhod tebe točno uže ne ponadobitsâ)

'with girl-oh look here, you write to me once more in neutral gender and you definitely won't need the exit any more)'

(17)
Мда... по-немецки "девушка" - Das Mädchen, это я знал.
Но блин, даже никогда не приходило в голову, понимаешь ли, что это средний род!
Становится понятным происхождение слова ***дéвушко*** (blog, 2007)

Mda... po-nemecki "devuška" - Das Mädchen, èto â znal.
No blin, daže nikogda ne prihodilo v golovu, ponimaeš' li, čto èto srednij rod!
Stanovitsâ ponâtnym proishoždenie slova ***dévuško***

'Well… 'the girl' is Das Mädchen in German, that I knew.
But jeez, I never even thought, you know, that it's *neuter gender*! The provenance of the word **devuško / girl-oh** becomes much more understandable.'

(18)
*Только что с Канзаки отожгли так, что чуть не померли со смеху :lol: Суть: я качала "Алхимика" с 27 по 43 эпизод, предки спят - ибо ночь, бабуся спит рядом - ибо мы в одной комнате <т.е. можети проснуться от шума> (...) Клинить нас начало после "бáбко"... :alles: и пошло-поехало... :zhosh: (...) 43 эпизóдо... и даунлóудо мáстеро сказало, что место нéто... надо дúско писáто... *че за язык нах?! :-D круче албанского!**[28] (blog, 2006; the post is entitled "night-oh language-oh")

*Tol'ko čto s Kanzaki otožgli tak, čto čut' ne pomerli so smehu :lol: Sut': â kačala "Alhimika" s 27 po 43 èpizod, predki spât - ibo noč', babusâ spit râdom - ibo my v odnoj komnate <t.e. možeti prosnut'sâ ot šuma> (...) Klinit' nas načalo posle "bábko"... :alles: i pošlo-poehalo... :zhosh: (...) 43 èpizódo... i daunlóudo mástero skazalo, čto mesto néto... nado dísko pisáto... *če za âzyk nah?! :-D kruče albanskogo!**

[28] :lol:, :alles: and :zhosh: denote various emoticons. In this example they are most probably incorrectly rendered HTML-codes, though intentional use of these codes instead of emoticons themselves cannot be ruled out.

'We've just really been on fire, Kanzaki and me, we almost died laughing :lol: To cut a long story short: I was downloading "The Alchemist", episodes 27 through 43, the 'rents are asleep because it's night, the grandma is sleeping next to us because we're in the same room <meaning she could wake up if we make too much noise> (…) We started to lose it after 'grandma-oh' … :alles: and so it went… :zhosh: (…) 43rd episode-oh… and download-oh master-oh said, that there is no-oh space… you should write-oh to the disk-oh… *what the fuck is this language?! :-D it's cooler than Olbanian[29]!*'

COMMENTARY: the language game the author and her friend are playing goes even further than the conversion-to-neuter seen in other examples. Here, *-o* is just being added to all the words: the process begins with nouns (*bábko* instead of *bábka* 'grandma'; *èpizódo* instead of *èpizód*) and then is expanded to other parts of speech (*skazalo* 'said-SG.N' instead of *skazal* 'said-SG.M', *pisato* instead of *pisat'* 'to write', *neto* instead of *net* 'no' etc.). As a result of this, words lose their declinability: thus, instead of *mesta* 'space-GEN.SG' we see *mesto* 'space-NOM.SG'.

(19)
А меня вот меня прикалывает все в среднем роде называть... Захотел недавно у девчонки очкú взять примерить и говорю: "Дáшко, дáй-ко мне примерить твоё очк... " (bash.org.ru, 2007)

A menâ vot menâ prikalyvaet vse v srednem rode nazyvat'... Zahotel nedavno u devčonki očkí vzât' primerit' i govorû: "Dáško, dáj-ko mne primerit' tvoe očk... "

'I have fun with calling everything in the neuter gender… Recently I wanted to try on one girl's glasses, so I tell her: "Dashk-oh, let-oh me try on your-oh glass…"'

COMMENTARY: The author's original intention had been to say *očkó* instead of *očkí* 'glasses', but then he must have realized that would sound rude (*očko* being a slang word for 'anus'), so he stopped abruptly.

EXPECTED STANDARD FORMS FOR 'Dashka, let me try on your glasses':

Дáшк-а	*дáй-ка*	*мне*
Dášk-a	*dáj-ka*	*mne*
Dashka-NOM.SG.D2	give:IMP-IMP	I:dat

[29] Olbanian language is another name for *âzyk padonkaff*, see 2.1.

примерить	*тво-и*	*очк-и́*
primerit'	*tvo-i*	*očk-í*
try_on:INF	your-ACC.PL	glasses-ACC.PL

USED FORMS:

Да́шк-о	*да́й-ко*	*мне*
Dášk-o	*dáj-ko*	*mne*
Dashka-NOM.SG.D1	give:IMP-IMP	I:dat

примерить	*тво-ё*	*очк[-ó]*
primerit'	*tvo-ё*	*očk[-ó]*
try_on:INF	your-ACC.SG.N	glasses-ACC.SG.D1

Comments in examples 16-17 are typical: in 16, the speaker expresses her attitude towards the innovation (probably jokingly, as the smiley indicates); in 17 the author is engaged in "folk linguistics" speculating about the origin of the innovation (probably also not altogether seriously). In example 18 the girl and her friend invent the game "transfer everything into the neuter gender" — on their own, as they seem to think. It is quite possible, though, that the already popular trend has helped this invention along.

Example 19 is especially interesting: the author claims that the new neuter also covers spoken language. Moreover, the author thinks that he is pronouncing the word *Да́шко / Dáško* in the neuter gender, even though in standard pronunciation it could not be distinguished from *Да́шка / Dáška*. This might be an illusion of difference between unstressed *-a* and *–o*, typical for a lay speaker misguided by orthography; on the other hand, the author could get engrossed in the game and indeed does not reduce the *-o* strongly enough. There is anecdotal evidence that this phenomenon is sometimes observed in spoken language (*kóšk[o], kís[o], Pólink[o]* instead of standard *kóšk[a]* 'cat', *kís[a]* 'pussycat', *Polínk[a]*).

Let us try to evaluate the status of this innovation in usage, applying not just the metalinguistic arguments of the speakers, but other examples as well.

It is obvious that the *-о* spelling has playful origins and is perceived accordingly. It seems, though, that as time went by it lost a great deal of its novelty, stopped being associated with a mutiny against the norm and became generally less conspicuous, transforming into something resembling the sort of half-jocular irregularity often used (and quite acceptable) in informal communication. Agreement in the neuter gender is either part of the same joke (amplifying it) or just automatic: if there's an *-о* ending, this must be neuter. Intuitively it seems that in most cases the agreement is automatic, but it is difficult, if not impossible, to prove this. In any case, the orthographic neuter becomes habitual, which makes it possible, for example, to perform the following playful reframing:

(20)
xxx: Любо-дóрого))
yyy: Кáтько - дéшево... (bash.org.ru [chat], 2009)

xxx: Lûbo-dórogo))
yyy: Kát'ko - déševo...

COMMENTARY: *любо-дóрого/lûbo-dórogo* is an idiomatic expression meaning 'it is a real pleasure to look at smth.', and that's what *xxx* is saying about something. However, the second interlocutor, yyy, playfully reanalyzes it as a sentence *Лю́ба — дóрого / Лю́ba — dórogo* with the female name *Лю́ба / Лю́ba* converted to neuter *Лю́бо / Лю́bo*. The sentence means 'Lûba-oh is expensive' or 'Lûba-oh is an expensive girl'. This reanalysis becomes clear from yyy's response: *Кáтько — дéшево* 'Kat′ka-oh is cheap'/'Kat′ka-oh is a cheap girl' (we can guess that Kat′ka is probably xxx's name).

First cue according to yyy's reanalysis ('Lûba is expensive'):

ALLEGED STANDARD FORMS	
Лю́б-а	*дóрог-о*
Лю́b-a	*dórog-o*
Lûba-NOM.SG.D2	expensive-PRED[30]

[30] Predicative is sometimes considered to be a separate part of speech, sometimes a function of short-form adjectives. It always has the same form as a neuter short-form adjective and an adverb.

USED FORMS:

Лю́б-о	*до́рог-о*
Лю́b-o	*dórog-o*
Lûba-NOM.SG.D1	expensive-PRED

Second cue (‘Kat′ka(-oh) is cheap’):

EXPECTED STANDARD FORMS:

Ка́тьк-а	*де́шев-о*
Ка́t'k-a	*déšev-o*
Kat'ka-NOM.SG.D2	cheap-PRED

USED FORMS:

Ка́тьк-о	*де́шево*
Ка́t'k-o	*déšev-o*
Kat'ka-NOM.SG.D1	cheap-PRED

Ten years earlier, in 1999, it would have been much more difficult to understand this joke, and the probability of its occurrence would have been much lower.

The orthographic neuter is sometimes used by very educated speakers, not necessarily in completely informal situations:

(21)
второе заме́тко сильно невнятное (mailing list, 2009; written by a respectable science journalist aged no less than 50)

vtoroe zamétko sil'no nevnâtnoe

‘The second article-oh is really inarticulate’

EXPECTED STANDARD FORMS FOR: ‘the second article is really inarticulate’

втор-ая	*заме́тк-а*	*сильно*	*невнятн-ая*
vtor-aâ	*zamétk-a*	*sil'no*	*nevnâtn-aâ*
second-NOM.SG.F	article-NOM.SG.D2	strongly	inarticulate-NOM.SG.F

USED FORMS:

втор-ое	*заме́тк-о*	*сильно*	*невнятн-ое*
vtor-oe	*zamétk-o*	*sil'no*	*nevnâtn-oe*
second-NOM.SG.N	article-NOM.SG.D1	strongly	inarticulate-NOM.SG.N

This shows that the transition is not just mass-scale and productive, but also habitual and acceptable for a certain stratum of speakers.

2.6. Beyond the “colloquial” Internet

Does the transition to neuter gender occur outside blogs, forums, chat rooms, personal correspondence and other media of informal communication? As section 4 will show, it happens very rarely.

In the existing examples, as opposed to examples from the “colloquial” Internet, the transition is usually marked by obvious semantic connotations: it stresses the asexuality, inanimate nature or low value of the creature or object:

(22)
Почин Совета Европы отказаться от слов "отец" и "мать", заменив их универсальным "родитель", кажется мне поистине великим, но недостаточно радикальным (...) надо придумать для обозначения человеческой особи слово среднего рода (...) Само собой, любая оценочная терминология упраздняется автоматически. "Красивая девушка" – пардон, "красивое дéвушко" – унижает некрасивое, которое стоит рядом, слышит это и может быть травмировано. "Хорошее фúльмо" – но тó человéк, которое сняло нехорошее, может обидеться и покончить с собой (...) (“Грешный мой язык”, Дмитрий Быков, “Труд” № 167, 09.09.2010)

Počin Soveta Evropy otkazat’sâ ot slov "otec" i "mat’", zameniv ih universal’nym "roditel’", kažetsâ mne poistine velikim, no nedostatočno radikal’nym (...) nado pridumat’ dlâ oboznačeniâ čelovečeskoj osobi slovo srednego roda (...) Samo soboj, lûbaâ ocenočnaâ terminologiâ uprazdnâetsâ avtomatičeski. "Krasivaâ devuška" – pardon, "krasivoe dévuško" – unižaet nekrasivoe, kotoroe stoit râdom, slyšit èto i možet byt’ travmirovano. "Horošee fil’mo" – no tó čelovék, kotoroe snâlo nehorošee, možet obidet’sâ i pokončit’ s soboj (…)(“Grešnyj moj âzyk”, Dmitrij Bykov, *Trud* 167, September 9, 2010)

‘The Council of Europe initiative to drop the words “father” and “mother” in favor of the universal “parent” is in my opinion a huge but insufficiently radical step. (…) One should invent a neuter-gender word to describe a human being (…) Obviously, any evaluative terms are automatically banished. “A beautiful girl” – sorry, “beautiful-oh girl-oh” – humiliates the ugly-oh [one] standing-oh next to her, who-oh can be traumatized-oh as a consequence. “A good-oh movie-oh” — but another-oh person who-oh shot-oh the bad-oh [one] could get offended and commit suicide (…)’

EXPECTED STANDARD FORMS FOR ‘[the expression] “a beautiful girl” humiliates the ugly [one] who is standing next to [her] and can be traumatized’**:**

красив-ая	*дéвушк-а*	*унижа-ет*		
krasiv-aâ	*dévušk-a*	*uniža-et*		
beautiful-NOM.SG.F	girl-NOM.SG.D2	humiliate-PRS.3SG		
некрасив-ую	*котор-ая*	*сто-ит*	*рядом*	
nekrasiv-uû	*kotor-aâ*	*sto-it*	*râdom*	
ugly-ACC.SG.F	which-NOM.SG.F	stand-PRS.3SG	next_to	
и	*мож-ет*	*быть*	*травмирован-а*	
i	*mož-et*	*byt'*	*travmirovan-a*	
and	can-PRS.3SG	be.INF	traumatize:PTCP:PASS-SG.F	

USED FORMS:

красив-ое	*дéвушк-о*	*унижа-ет*		
krasiv-oe	*dévušk-o*	*uniža-et*		
beautiful-NOM.SG.N	girl-NOM.SG.D1	humiliate-PRS.3SG		
некрасив-ое	*котор-ое*	*сто-ит*	*рядом*	
nekrasiv-oe	*kotor-oe*	*sto-it*	*râdom*	
ugly-ACC.SG.N	which-NOM.SG.N	stand-PRS.3SG	next_to	
и	*мож-ет*	*быть*	*травмирован-о*	
i	*mož-e*	*byt'*	*travmirovan-o*	
and	can-PRS.3SG	be.INF	traumatize:PTCP:PASS-SG.N	

EXPECTED STANDARD FORMS FOR ‘“A good movie” — but another person who filmed a bad [one]...’:

хорош-ий	*фи́льм-ø*	*но*	*тóт*	*человéк-ø*
horoš-ij	*fíl'm-ø*	*no*	*tót*	*čelovék-ø*

good-NOM.SG.M movie-NOM.SG.D1 but that:SG.M person-NOM.SG.D1

котор-ый *снял-ø* *нехорош-ий*
kotor-yj *snâl-ø* *nehoroš-ij*
which-NOM.SG.M film:PST-SG.M bad-NOM.SG.M

USED FORMS:
хорош-ее *фи́льм-о* *но* *то* *челове́к-ø*
horoš-ee *fil'm-o* *no* *to* *čelovék-ø*
good-NOM.SG.N movie-NOM.SG.D1 but that:SG.N person-NOM.SG.D1

котор-ое *снял-о* *нехорош-ее*
kotor-oe *snâl-o* *nehoroš-ee*
which-NOM.SG.N film:PST-SG.N bad-NOM.SG.N

Zubova (2010, 20; 2000) provides a number of examples from poetry: *ЧЬЕ ОНО ЛЮБО́ВНИЦО? / Č'E ONO LJUBÓVNICO?* / 'WHOSE-OH LOVER-OH (IS) THIS-OH?' (V. Sosnora, 1976–1978); *певчее зе́во свое / pevčee zévo svoe* / 'one's-oh singing-oh gorge-oh' (Yu. Kublanovskij, 1986), where the transition is used as a poetic device.

Sometimes in the literature and press of the 21st century, the transition is used without any obvious semantic load. It was used in this way earlier, too: Norman (2006, 168) quotes V. Levi's («Искусство быть другим» / "Iskusstvo byt' drugim", 1980): *ваше арбу́зо, между прочим, треснутое / vaše arbúzo, meždu pročim, tresnutoe* / 'your watermelon-oh, by the way, is cracked-oh'. In the earliest found occurrence, too, there is no obvious semantic load.

(23)
...стены беседки испещрены всевозможными надписями, как-то: «Никогда не забуду ночи на 8 августа»; «дунь-душа душенька»; «доктора скоты»; «я люблю красивое де́вушко» и некоторыми другими... (Иван Щеглов, "Дачный муж", 1888)[31]

...steny besedki ispeŝreny vsevozmožnymi nadpisâmi, kak-to: «Nikogda ne zabudu noči na 8 avgusta»; «dun'-duša dušen'ka»; «doktora skoty»; «â

[31] The example is transcribed in post-1918 spelling.

lûblû krasivoe dévuško» i nekotorymi drugimi... (Ivan Ŝeglov, "Dačnyj muž", 1888)

'The gazebo's walls are peppered with various inscriptions, such as: "I'll never forget the night of August 8"; "doodle-dolly darling"; "doctors are pigs"; "I love a beautiful-oh girl-oh" and some others…'

EXPECTED STANDARD FORMS FOR 'I love a beautiful girl':

я	*любл-ю*	*красив-ую*	*дéвушк-у*
â	*lûbl-û*	*krasiv-uû*	*dévušk-u*
I	love-PRS.1SG	beautiful-ACC.SG.F	girl-ACC.SG.D2

USED FORMS:

я	*любл-ю*	*красив-ое*	*дéвушк-о*
â	*lûbl-û*	*krasiv-oe*	*dévušk-o*
I	love-PRS.1SG	beautiful-ACC.SG.N	girl-acc.SG.D1

One is forced to ask: if, in fiction, the transition into the neutral gender is found as long ago as the 19th century, does it make sense to say that the innovation was born thanks to the Internet?

2.7. The role of the Internet in the origin of the innovation

In examples from fiction and press in the 19th and 20th centuries, the transition to the neuter gender is occasional. Sometimes, it has obvious connotations, sometimes they are less obvious, but the general function of the transition can be described as "artistic device". It looks strange and therefore stands out. In the 21st-century Internet, as discussed in 2.5, the transition becomes a productive model, its strangeness fades, while its popularity grows.

The scope for this transition had existed in the Russian language system before, but only on the Internet has its use become mass-scale, allowing us to speak of a grammatical innovation. The innovation so far remains typical for the "colloquial" registers of the Internet, but sometimes ventures beyond their boundaries.

Let us list the features of open colloquial registers of the Internet, such as blogs, forums, and multi-user chat rooms, which were instrumental in transforming a rare

wordplay into a productive model. First, the Internet provides a written channel of communication: in standard oral speech, *дéвушка / dévuška* and *дéвушко / dévuško* are indistinguishable.[32] Second, in those registers deviations from codified standards are acceptable and even encouraged[33]: in newspapers, books, documents and other traditional written registers this is virtually impossible (especially if the text is reviewed by professional editors before publication). Third, these registers differ from traditional written communication since the message is available not just to the addresser and addressee (and possibly their close circle), but to *anyone*. The structure of information exchange (and, moreover, the structure of social network) changes: anyone can easily write a text that would not be edited by anyone, but would be potentially available to millions of users for an (again potentially) unlimited time. This feature of the channel makes the rapid spread of innovations possible.

As a consequence, any innovation has a chance of not disappearing immediately after its appearance, but becoming noticeable and popular extremely soon (in linguistic terms), in a matter of years.

It is to the three factors listed that we owe the orthographic neuter. The written channel and special relationship with the standard produce the *âzyk padonkaff* that quickly gains popularity. The fashion for the *âzyk padonkaff* later passes, but one of its features produces an innovation that penetrates the grammatical system of colloquial Russian on the Internet.

[32] Now that the innovation is widespread, one can imagine artificial oh-stressing, intended to imitate the innovation in oral speech (see section 2.5), but it is hard to imagine people making this effort before the transition to the neuter gender has become popular.

[33] The reasons for such an attitude towards the norm are not completely clear and should be studied. One might assume that the high speed of communication (including the speed of message exchange, i. e. interactivity) and universal access to the messages (see below in the text) ought to figure among those reasons.

3. Role of factors other than the Internet

In the course of discussing this article with my colleagues, I faced critical remarks concerned mostly with my treatment of factors influencing the process described; in particular, over-emphasizing the role of the Internet. This section analyzes such remarks, as well as a number of alternative interpretations of the innovation and its origin as offered by Zubova (2010).

It is extremely unlikely that mass transition to neuter gender occurred under the influence of other languages: Ukrainian nominative *-о* endings (*дядько / djad'ko /* 'uncle') or vocative in some declensions (*мамо / mamo /* 'mommy') or the German *–chen* suffix (as the author of example 17 seems to think). There is no evidence to support this; moreover, neither Ukrainian nor German is popular enough among Russian speakers to play a significant role.

Krongauz (2008) and Zubova (2010, 17) write that forms ending in *-о* spread as an analogy to the *йа кривéтко / ja krivétko* cliché / 'I am a shrimp-oh' (an internet meme, once extremely popular). This is incorrect: *кривéтко / krivétko* and its variants appear in 2007 (see Fig. 1), while forms ending in *–о* appear several years earlier; by 2006 they are already popular. This creates the background which can foster the growth of *кривéтко / krivétko* (which possibly accelerates further transitions to the neuter).

Zubova (2010, 18–19, 22) lists a number of parallels with the history of the Russian language. Indeed, a new transition to the neuter immediately evokes a number of historical associations in any linguist. However, the fact that the new forms are superficially similar to some forms that existed at some earlier stages of the language does not per se mean that these earlier forms played a role in the new transition. It is also necessary to demonstrate how such influence could be exercised. In some cases the influence mechanism is evident. Thus, for example, the neuter gender of the word *колéно / koléno /* 'knee' could contribute to the transition of the diminutive *колéнка / kolénka /* 'knee' back into the neuter; the tendency of words with diminutive-

evaluative suffixes *-ишко / -iško* to agree in the neuter[34] never disappeared from spoken language, which certainly facilitates renewed activation of the same model. In some cases, though, the influence mechanism is not just unsupported by evidence, but seems virtually impossible: the vocative ending in *-о* (*же́но* / *žéno* / 'wife') had long disappeared from the Russian language, and it is not clear how it could facilitate the spread of neuter words denoting people in the 21st century. The potential mechanism of the influence of words like *ча́до / čádo* / 'child' and *дитя́ / ditá́* / 'infant' on oh-words *ребёнко / rebёnko* (instead of *ребёнок / rebёnok* /'child'), *вну́чко / vnúčko* (instead of *вну́чка / vnúčka* / 'granddaughter') and *племя́нницо / plemá́nnico* (instead of *племя́нница* / *plemá́nnica* / 'niece') also remains unclear.

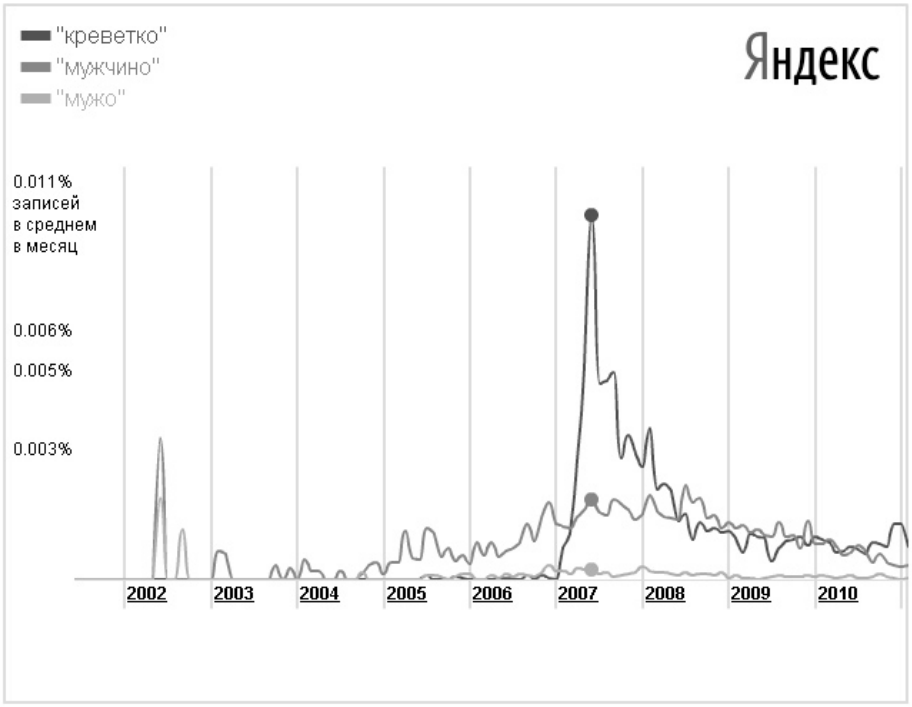

Fig. 1. Relative frequency of occurrences *кре́ветко / krévetko* / 'shrimp-oh' *, мужчи́но / mužčíno* / 'man-oh' and *му́жо / múžo* / 'husband-oh' in Russian-language blogs according to Yandex's "Pulse of the Blogosphere" service: the line with the large peak in 2007 is *krevétko*; the line with a small peak in 2002 is *mužčíno*; the line mostly oscillating around zero is *mužo*. The variants *krivétko, krevédko* and *krivédko* behave more or less the same as *krevétko* while not being as frequent; *mužčíno* and *múžo* were selected as the most frequent representatives of the two classes of nouns reviewed in section 4.

[34] Originally they all belonged to the neuter gender (Nejdâeva 1985).

Any significant influence of hyperonyms *живóтное / živótnoe /* ‘animal’*, насекóмое / nasekómoe /* ‘insect’*, млекопитáющее / mlekopitáûŝee /* ‘mammal’ on the transition of words denoting people into the neuter, suggested by Zubova, also seems unlikely: hyperonyms do not seem to be so active in the speakers’ mind as to produce playful forms of the *дéвушко / dévuško /* ‘girl-oh’ and *мужчúно / mužčíno /* ‘man-oh’ type.

Zubova’s hypothesis (Zubova 2010, 23), that the transition into the neuter is largely based on the hypothesized original semantics of the gender described in general terms by the third-century Christian philosopher Ammonius of Alexandria, does not seem likely, either. As noted above, the transition, once it became mass-scale, largely lost its semantic significance. In some cases it is certainly present, especially in words denoting people. It is, however, quite often impossible to point out what exactly the transition to the neuter adds to the word’s meaning; see examples 2 (birthday-oh), 4 (girl-oh), 5 (service cap-oh), 6 (book-oh), 8 (bitch-oh), 10 (car-oh), 12 (day-oh), 18, and 19. Zubova (2010, 21) suggests that the new neuter can be used to indicate that a person does not really fit his or her gender role. This explanation, however, clearly does not work in example 11: it is very unlikely that the ‘asexual’ nature of the child’s father is stressed (on the contrary, this is a typical folklore portrait of a male software engineer). A pejorative attitude, often associated with the neuter, is also unlikely: the wider context shows that the speaker (the child’s mother) treats the “admin-oh” with sympathy. See also the following example:

(24)
а вообще готовишь вкусно?? Я как мужчúнко спрашиваю)))))
(bash.org.ru [chat], 2012)

a voobŝe gotoviš’ vkusno?? Â kak mužčínko spraševaû)))))

‘Are you a good cook generally? I’m asking as a man-oh))))))’

EXPECTED STANDARD FORM FOR ‘a (small) man’:

мужчúн-а		*мужчúн-к-а*
mužčín-a	OR	*mužčín-k-a*
man-NOM.SG.D2		man-DIMIN-NOM.SG.D2

USED FORM:
мужчи́н-к-о
mužčín-k-o
man-DIMINUTIVE-NOM.SG.D1

COMMENTARY: the diminutive form *mužčínka* is highly unlikely in this context (see below), so the more probable standard counterpart is *mužčína*. The new neuter, however, is often used together with the diminutive suffix *-k-*.

It is improbable that the speaker wants to stress pejorative self-image or his own asexuality, especially since the features in question are stereotypical for a "typical guy" (loves to eat, the cooking should be done by a woman), and the general communicative goal of the communication seems to be to flirt with the girl.

In other words, for any specific example it is possible to find an *ad hoc* description of semantic significance that the transition to the neuter gender would entail. However, such descriptions are of little value unless they explain many (ideally all) examples.

In light of this, the following remark by Zubova is very important (Zubova 2010, 21): "The grammatical play activating the neuter gender implements the human need to subdue the pomposity of the statement, to make speech more intimate". This observation seems very accurate: Internet users indeed often try to subdue pomposity, and the neuter gender provides a handy tool for that. It does not necessarily lower the evaluation of a specific object, but it makes the statement (or a part thereof) less serious, more ironic, in accordance with the style of the register as a whole. It is possible that this factor makes the transition of the final unstressed *-a* to *-o* one of the most popular and noticeable traces of the *âzyk padonkaff*.

Some of the listed factors, therefore, could indeed influence the spread of the innovation (most prominently the need to subdue the pomposity of a statement). But even taken together they do not create sufficient conditions to launch a transition. This is demonstrated by the time when the innovation occurred: if the Internet had not been a necessary condition, why did the mass-scale transition not start sooner?

All factors listed in this section (except the word *krivetko*) had appeared long before the 21st century.

4. Quantitative verification

Some of the assertions made in this article may be verified quantitatively. If it is true that the transition to the neuter is a mass-scale event, then many examples of that kind should exist. If it is true that the transition mostly covers nouns ending in unstressed *-a*, respective forms for them should be encountered more frequently.[35] If it is true that the transition occurs in the "colloquial" Internet in the 21st century, it should occur less frequently in other registers and before the set date.

The following calculations were performed to check these hypotheses. Ten nouns ending in unstressed *-a* (class A) were selected for comparison, as well as 10 nouns with other endings (class B). The most frequent nouns naturally occurring in live oral speech[36], according to a frequency dictionary (Lâshevskaâ, Šarov 2008), were selected. Nouns with stem ending in palatalized consonant *(мать / mat' /* 'mother'*, история / istoriâ /* 'story' and *чай / čaj /* 'tea') were excluded for simplification, so that all oh-words would end in *-o*, not in *-e/-ë*. Among other parameters, only animacy was controlled: both samples have five animate and five inanimate nouns.

The calculations were performed in different sources using different methods (see Tables 1a and 1b). In three subcorpora (main, newspaper, dialect) of the National Corpus of the Russian Language (column NCRL), the search query was the oh-word

[35] This hypothesis is partially supported by examples 18 and 22. It may be no accident that the "night-oh language-oh" in example 19 starts off with *бáбко / bábko /* 'grandmother-oh' (and not, for example, with *мáстеро / mástero /* 'máster-oh'). In example 23 there is a remarkable phrase *мó человéк, которое / tó čelovék, kotoroe /* 'the-oh man who-oh'. The form *дéвушко / dévuško /* 'girl-oh' apparently does not cause any difficulty for the author, but the form *человéко / čelovéko /* 'man-oh', conspicuous by its absence in agreement with the neuter gender, does seem to do so.

[36] The vocabulary of this register can be assumed to be close to the vocabulary of registers studied in this article.

ending in *-o*. The same query was used to search the Sexnarod forum, one of Runet's largest forums, using the Google search page (column 'Forum' in the table). Among all found pages of the forum, examples where oh-word was agreed in the neuter gender were selected manually; their number is indicated in subcolumn *n*. Subcolumn *N* shows how many times the corresponding noun is used on the forum in standard form, without the *-o* ending (these results were not reviewed manually). Subcolumn *w* shows the weighted frequency of oh-word *n/N*. This result is used as the main indicator of how frequently the given noun is being converted into neuter.

In other sources the straightforward search for oh-words produced too much noise, and it was refined. Using the bigram search of aot.ru group,[37] the most frequent left and right one-word contexts for each noun were determined. Five adjectives were taken from the left context, five verbs from the right context. For each noun 11 search queries were run: oh-word and agreed adjective (*настоящее мужчи́но / nastojaŝee mužčína* / 'real-oh man-oh'); verb (*мужчи́но было / mužčína bylo* / 'the man-oh was-oh'), or the pronoun *which* (*мужчи́но которое / mužčína kotoroe* / 'the man-oh which-oh'). The results were reviewed manually. This method was used to search the mass media database "Integrum"[38] (the column "Integrum), the Google Books service ("Books") and the Yandex Blogs service ("Blogs"). For the blogs, the subcolumn *n* shows the total of all 11 queries, the subcolumn *f* the frequency of the noun according to the Lyashevskaya and Sharov dictionary (the Yandex search engine does not make it possible to define the frequency of a given word in blogs with sufficient precision), the subcolumn *w* shows the weighted frequency of the oh-word *n/f*.

It is evident that all three hypotheses are confirmed. The phenomenon is mass-scale: for 20 random nouns[39] there are 210 examples of transition, on a forum (a major one,

[37] http://aot.ru/demo/bigrams.html

[38] http://www.integrumworld.com/

[39] Strictly speaking, not truly random, but rather the most frequently selected with regard to being animated, but it is important to note that they were selected irrespective of any data on their transition to the neuter gender.

but just one forum) there are 12. The phenomenon is virtually limited to the "colloquial" Internet: there is one example found outside it. For class A the transition occurs more frequently than for class B, and the difference is statistically significant. For the blogs, p=0.03 (two-sided Welch t-test), the effect size (Cohen's d) is 1.11 (usually the effect is considered to be large at $d > 0.80$); for the forum, p=0.05, d=1.02 respectively; in both cases weighted frequencies are compared.

Anim.	Word	Blogs			Forum			NCRL	Integrum	Books
		n	f	w	n	N	w			
anim.	*мужчи́но* *mužčíno* 'man-oh'	44	99,11	0,444	6	854	0,007	0	0	0
anim.	*кры́со* *krýso* 'rat-oh'	16	33,98	0,471	2	179	0,011	0	0	0
anim.	*мама́шо* *mamášo* 'mother-oh'	2	70,79	0,028	1	415	0,002	0	0	0
anim.	*ча́йко* *čájko* 'seagull-oh'	3	16,99	0,177	0	103	0,000	0	0	0
anim.	*ста́рушко* *stáruško* 'old_woman-oh'	22	14,16	1,554	2	175	0,011	0	0	0
inanim.	*середи́но* *seredíno* 'middle-oh'	7	39,64	0,177	0	246	0,000	0	0	0
inanim.	*практи́ко* *praktíko* 'practice-oh'	1	124,60	0,008	0	509	0,000	0	0	0
inanim.	*кры́шо* *krýšo* 'roof-oh'	30	45,31	0,662	0	288	0,000	0	0	0
inanim.	*карти́но* *kartíno* 'picture-oh'	19	73,63	0,258	1	474	0,002	0	0	0
inanim.	*нау́ко* *nauko* 'science-oh'	5	42,48	0,118	0	354	0,000	0	0	0
Total for class A:		**149**		**3,896**	**12**		**0,034**	**0**	**0**	**0**

Table 1a. Relative frequencies of transition to the neuter for nouns ending in unstressed -*a*.

Anim.	Word	Blogs			Forum			NCRL	Integrum	Books
		n	f	w	n	N	w			
anim.	*му́жо* *múžo* 'husband-oh'	36	688,11	0,052	0	845	0,000	0	0	0
anim.	*дире́кторо* *diréktoro* 'director-oh'	1	59,47	0,017	0	336	0,000	0	0	0
anim.	*ко́то* *kóto* 'tomcat-oh'	10	101,94	0,098	0	775	0,000	0	0	0
anim.	*ма́стеро* *mástero* 'master-oh'	1	59,47	0,017	0	495	0,000	0	0	0
anim.	*чува́ко* *čuvákо* 'dude-oh'	0	59,47	0,000	0	357	0,000	0	0	0
inanim.	*сторонó* *storonó* 'side-oh'	3	294,5	0,010	0	373	0,000	0	0	0
inanim.	*магнитофóно* *magnitofóno* 'tape_recorder-oh'	0	82,12	0,000	0	351	0,000	0	0	0
inanim.	*вопро́со* *vopróso* 'question-oh'	6	161,41	0,037	0	845	0,000	0	0	0
inanim.	*гри́бо* *gríbo* 'mushroom-oh'	0	141,59	0,000	0	93	0,000	0	0	0
inanim.	*моме́нто* *momén̄to* 'moment-oh'	4	90,62	0,044	0	823	0,000	0	0	1
Total for class B:		**61**		**0,276**	**0**		**0,000**	**0**	**0**	**1**

Table 1b. Relative frequencies of transition to the neuter for nouns not ending in unstressed -a[40]

[40] For class A, the original form can be reconstructed by changing the final *o* into *a*, and the same applies for the word *storoná* from class B. In the remaining nine cases, the final *o* should simply be deleted.

5. Conclusion

The hypothesis about the influence of the communication channel on language phenomena is often proposed in connection with the alleged influence of writing on the development of language (see the review in Wray and Grace 2007, 557–559). Unfortunately, it is difficult to verify this hypothesis for this channel: it is hard to ascertain what the languages were like before the spread of writing and literacy. The Internet is much more accessible for this type of research, and its potential influence on language is probably as powerful as that of writing. My main goal was to show one example of such influence.

It is difficult to predict the future of this specific example, "the orthographic neuter". It may remain a widespread, playful irregularity, used without investing too much meaning in it; it may also lose popularity and revert to the 'dormant' state.

Acknowledgements

I am grateful to Martin Paulsen, who read an earlier draft of this article and made a number of insightful remarks, to everyone with whom I have discussed the phenomena in question, especially Aleksandr Piperski, Vladimir Plungân and Lûdmila Zubova, and to Viktor Sonkin, who translated the article into English.

References

Biber, Douglas. 1994. "An Analytical Framework for Register Studies." In *Sociolinguistic perspectives on Register*, edited by Biber Douglas and Edward Finegan, 44–56. New York: OUP.

Hristova, Daniela. 2011. "Velikij i mogučij olbanskij âzyk: The Russian Internet and the Russian Language." *Russian Language Journal* 61: 141–160.

Il′ina, N. E. 1996. “Rost analitizma v morfologii.” [Analytic trends in Russian morphology] In *Russkij âzyk konca XX stoletiâ (1985–1995)* [Russian language at the end of 20th century (1985-1995)], edited by Elena Zemskaâ, 326–343. Moscow: Âzyki russkoj kul′tury.

Krongauz, Maksim. 2008. “Utomlennye gramotoj.” [Wearied by literacy] *Novyj mir* 5: 153–159. Accessed at http://magazines.russ.ru/novyi_mi/2008/5/kr11.html

Lâševskaâ, Olga; and Šarov, Sergej. 2009. *Novyj častotnyj slovar′ russkoj leksiki* [New frequency dictionary of Russian words]. Moscow: Azbukovnik. http://dict.ruslang.ru/freq.php

Levin, Maurice. 1978. *Russian declension and conjugation: a structural description with exercises*. Columbus: Slavica Publishers.

Nejdaeva, Svetlana. 1985. “Istoriâ grammatičeskogo razvitiâ suŝestvitel′nyh sub″ektivnoj ocenki.” [The history of the grammatical development of the subjective evaluation nouns] PhD diss., Alma-Ata.

Norman, Boris. 2006. “Lingvopsihologičeskie aspekty grammatičeskoj kategorii roda.” [Psycholinguistic facets of the grammatical category of gender] Russian Linguistics 30(2): 153–174.

Tagliamonte, Sali, and Denis, Derek. 2008. “Linguistic ruin? LOL! Instant messaging and teen language.” *American Speech* 83(1): 3–34.

Wray, Alison and George Grace. 2007. “The consequences of talking to strangers: evolutionary corollaries of socio-linguistic influences on linguistic form.” *Lingua* 117 (3): 543–578.

Zaliznâk, Andrej. *Grammatičeskij slovar′ russkogo âzyka. Slovoizmenenie* [Grammar dictionary of the Russian language. Inflection]. M., 1977.

Zubova, Lûdmila. 2000. “Kategoriâ roda i lingvističeskij èksperiment v sovremennoj russkoj poèzii.” [The gender category and the language experiment in modern

Russian poetry] In *Problemy funkcional'noj grammatiki: Kategorii morfologii i sintaksisa v vyskazyvanii* [Problems of functional grammar: morphological and syntactic categories in the utterance], edited by A. Bondarko and S. Šubik, 194–210. St Petersburg: Nauka. Accessed at http://www.levin.rinet.ru/ABOUT/zubova1.html, January 26, 2013.

Zubova, Lûdmila. 2010. "Ironičeskaâ grammatika: srednij rod v igrovoj neologii." [Ironic grammar: neuter gender in playful neology] *Voprosy âzykoznaniâ* 6: 16–25.

Zvereva, Vera. 2009. "«Âzyk padonkaf»: Diskussii pol′zovatelej Runeta." [The *padonki* language: discussions of Runet users] In *From Poets to Padonki. Linguistic authority & norm negotiation in modern Russian culture*, edited by Ingunn Lunde and Martin Paulsen, 49–79. Bergen.

Appendix 1. Brief description of relevant features of the Russian language

A1.1 Morphology

There are three genders in Russian: masculine, feminine and neuter. Generally speaking, gender is not semantically motivated, apart from most nouns for male and female beings. Gender is not intrinsically marked in the noun itself, though the correlations between gender and nom. sing. ending are quite strong. Thus, the ending [a] (orthographically *a* or *â*) is most typical for feminine nouns, zero ending for masculine nouns (the last sound of the word in this case is the final consonant of the stem), and the endings [o] (orthographically *o* or *ë*) and [e] for neuter gender (more on neuter gender and *o*-endings below).

Adjectives agree in gender with nouns (and personal pronouns). This is true both of long-form adjectives (in nom.sing.masc. the possible endings include *-yj, -ij, -oj;* in nom.sing.fem. — *aâ*, in nom.sing.neut. — *-oe, -ee*) and short-form adjectives (in nom.sing.masc. the ending is *-ø,* in fem. *-a*, in neut. *-o*). Other agreeing parts of speech include verbs (only in past tense singular; the set of endings is the same) and

anaphoric pronouns in the third person singular (the set of endings is the same). Gender, in a complicated fashion, also influences the case declension of nouns.

There are three main noun declensions, which can be given the following conventional labels: ’nonfeminine’ (masculine and neuter nouns ending in nom. in [o] or [e], and masculine nouns with zero ending); ’nonneuter’ (masculine and feminine nouns ending in nom. in [a]); ’feminine’ (feminine nouns ending in nom. in palatalized consonant). Three-way division is only relevant for the singular number; in the plural, almost all case endings are the same for the three declensions.

Indeclinable nouns do not change in different numbers and cases; the vast majority of such nouns belong to the neuter gender.

The innovation under review can modify the nouns’ declension; moreover, the same word can change declensions within one example. Declension thus becomes a de facto inflexional category and will be noted in glosses accordingly (d1 for nonfeminine, d2 for nonneuter, d3 for feminine, d0 for indeclinable).

Adjectives and pronouns are declined differently depending on their gender, but in the plural case endings are the same for all three genders. It is also important to note that, in the singular, adjectives of masculine and neuter gender only differ in nom. and acc., being the same in all other cases.

For a more detailed and precise description of Russian morphology, see Levin (1978).

A1.2. Phonology and orthography

Reduction of vowels and the phonemic principle of Russian orthography are also relevant to this article. The vowels [a] and [o] in standard Russian are only distinguishable in stressed position; in other positions they sound the same. Specific phonetic implementation in unstressed positions depends on the vowel’s position

within a word and neighboring sounds, but in many cases it sounds more like [a]. Russian orthography, though, is based on the phonemic principle (sometimes called 'morphological'): "write not the sound you hear, but the underlying phoneme"; as a result, hearing the unstressed [a], one quite often has to write *o*. This often leads to spelling mistakes: incorrect substitution of *a* instead of *o* is widespread ("as one hears"), as well as *o* instead of *a* (hypercorrection).

IV

4. The written turn and its role in actuating language change

1. Introduction

1.1. The actuation problem

More than 40 years ago, Weinreich, Labov and Herzog (1968, 102) formulated the so-called *actuation problem*: "Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times?". They referred to it as "perhaps the most basic" question facing students of language change.

Most sociolinguists acknowledge that the quest to solve the actuation problem does not have to be limited to internal factors. As Weinreich, Labov and Herzog (1968, 186) themselves note, stimuli and constraints that affect the process of language change can be found in both society and the structure of language. Well-recognized external causes of language change include, for instance, changes in social structure or language contact (see Cheshire, Agder and Fox 2012 for an analysis of how the combination of internal and external causes actuates a morphosyntactic change in English).

The general question this dissertation addresses is the role of computer-mediated communication (CMC) in language change. This chapter discusses some aspects of the potential role of CMC in the actuation of language change.

1.2. CMC and language change

CMC has long been recognized as a factor in language change. Baron (1984) made one of the very first declarations of interest, Herring (2012) and Androutsopoulos (2011a) offer brief reviews of what has been done since then (Androutsopoulos,

however, warns against overestimating the role of CMC in language change). The interest in CMC has reinvigorated the long-standing interest in the more general question of the role of the communication channel (medium) in language change (Herring 2003).

To the best of my knowledge, however, there has been little explicit discussion of whether studies of language change in CMC allow us to arrive at a partial solution of the actuation problem.

One reason might be that the actuation problem lies within the field of "classic" sociolinguistics, and this field's questions and methods are not exactly the same as those in CMC studies. Androutsopoulos (2011b), for instance, discusses several of the limitations of a variationist sociolinguistics approach to CMC. First, the narrow understanding of a linguistic variable (on various understandings of linguistic variables see Wolfram 1991, Tagliamonte 2006, 70–98) excludes certain features, which can nonetheless be of importance. Second, quantitative-based sociolinguistics excludes features that are scarce in frequency, even though they are not necessarily marginal. Third, sociolinguistic analysis often relies on non-linguistic parameters (age, gender, region) that are to a large extent predefined by scholarly conventions. These categories, however, are not necessarily those that are most relevant to variation in the linguistic behavior observed in CMC.

These divergences between the two fields can be seen quite clearly in this chapter. The innovations that will be discussed are not linguistic variables stricto sensu. For most of them, it is difficult to estimate the frequency reliably, for some, that frequency is likely to be low. Further, I do not show any link with traditional social parameters like age and gender (instead, I claim that a parameter that facilitates their emergence is the communication channel). Finally, it is possible that these innovations will be short-lived, and some of the new patterns are very unlikely to replace the existing ones, and may at best co-exist with them.

Nevertheless, all examples in this chapter attest to the fact that a new pattern has emerged in the speakers' linguistic behavior. In each case, I explain which pattern is new, what is new about it and why the innovation is worth attention. Milroy and Milroy (1985, 345) claim that "[i]f we are to address the actuation problem ... we must break with tradition and maintain that it is not languages that innovate; it is speakers who innovate". In my examples, speakers do innovate in some way, and the important question is — why do they do it?

1.3. A Two-stage model of innovation and its importance for CMC

It is important to distinguish between the two processes involved in language change (Croft 2000, 4–5): the initial *innovation* (sometimes called *actuation*) and the subsequent *diffusion* (or *propagation*). The two processes can also be referred to as *mutation* and *selection* (Maslova 2008) or *altered replication* and *differential replication* (Croft 2000). I refer to the former as *speaker innovation* (after Milroy and Milroy 1985), to the latter as *diffusion*, and to the result of diffusion as *community innovation*. The word *innovation* is used to denote any new pattern in the language, either one that is represented by a single speaker innovation or one that has already diffused into the community. There is, of course, no strict boundary between the two types of innovation, rather a continuum. A more detailed model of change is possible (and probably desirable): for instance, one that also singles out the *actualization* process (Andersen 2001), but for the purposes of the current chapter the simple two-stage model is satisfactory.

One of the key claims of this chapter is that most of the influence that CMC exerts on language change occurs at the stage of diffusion. In other words, if something new has appeared in language because of the existence of CMC, it is most likely that that has happened because the stage of diffusion has been affected, not the stage of speaker innovation. For many novel patterns, the emergence of which is typically

associated with CMC, very similar speaker innovations can be found in "traditional" channels, long before the advent of CMC.

A textbook example of a CMC-born innovation are emoticons, created, according to the most widely held belief, in 1982 (Fahlman ?2002). Zimmer (2007), however, provides examples of features which are virtually identical to emoticons, both in form and function, and which date back to 1967, 1969 and 1887 (sic!). All these were independent one-speaker innovations that were not adopted by other members of the community and did not diffuse further. Thus, the use of emoticons went unactuated prior to 1982 not because they had not been invented previously (they had), but because there was no medium where they could easily spread. Text-based CMC, with its lack of non-verbal cues, limited set of characters, high informality and playfulness and other properties, provided such a medium. It also provided a multiplex social network where the innovation was able to spread and establish itself extremely quickly. In other words, if we focus on speaker innovations, post-1982 emoticons are not really new. As community innovations, however, they are.

It might be that CMC also creates new opportunities for speaker innovations, and that some fundamentally novel speaker innovations occur. For instance, in CMC, language change can be actuated by a typo. Consider leet speak, which produced neologisms such as *pwn* ('own') and *teh* ('the'), both the result of a typo (Ross 2006), or Russian *ncmo/psto* ('blogpost'), most likely of the same origin. Even in this case, however, the main effect of CMC occurs at the stage of diffusion: typos could have occurred (and most likely did) in traditional channels, but only in CMC did speakers pick them up and start using them intentionally.

Taking all of this into account, I investigate the forces that are at work at the diffusion stage: not the structure of the social network and its role in the spread of innovation, as many sociolinguistic studies do (see Paolillo 2001 about the application of the social network approach to CMC), but rather the selective forces that either promote the speaker innovation or prevent it from diffusing into the community.

To analyze these selective forces, we must abandon the uniformitarian principle (Labov 1972), the assumption that linguistic forces operating today are the same as those that operated in the past, or at least in the not-too-distant past (Labov 1994). On the contrary, these forces can change when social structure changes (Trudgill 2011, 167–169), or when important cultural and technological changes take place. This view allows us to address the actuation problem.

To return to our earlier example: emoticons were actuated several times prior to the advent of CMC, but the selective forces against them were strong, and they did not diffuse. The spread of CMC changed the selective forces, and the actuation in 1982 was successful.

The advent of CMC, a communication channel with a peculiar set of properties, launched several trends that influence which selective forces are at work and in which direction they operate. In this chapter, I focus on the example of one such trend, which I call "the written turn" (after Baron 2005). The written turn can briefly be described as follows: due to the increased use of written speech, speaker innovations which can easily diffuse in written speech, but not in oral speech, enjoy much higher chances of success.

1.4. The written turn

The fact that CMC changes the traditional relationship between written and oral speech, and cannot easily be classified as either, has been mentioned on innumerable occasions. Jones and Schieffelin (2009), Tagliamonte and Denis (2008), Dresner (2005), Zitzen and Stein (2004), Ko (1996), Yates (1996), Collot and Belmore (1996) and Ferrara, Brunner and Whittemore (1991) discuss, either some of the aspects of this complex relationship, or the manifestation of oral and written features in CMC, or the status of CMC: hybrid register or the medium in its own right, or many other aspects of "written" and "spoken" constituents of CMC. Androutsopoulos (2011b) emphasizes the need to rethink the methods of "classic" sociolinguistics (that were

designed with oral speech in mind) in order to study language change in the visual environment. Baron (2005) offers a brief overview of the "written turn" in modern language culture and the role of CMC in it.

Of importance to this chapter are the novel patterns of linguistic behavior that provide evidence about this written turn, the changing role of written speech, its increased autonomy and salience. Werry (1996, 58), examining data from IRC, identifies "an almost manic impulse to produce auditory and visual effects in writing, a straining to make written words simulate speech". In section 2, I analyze several patterns in Russian CMC that can be classified as the results of the opposite "impulses". In section 2.1, I focus on a written-speech pattern that is in *direct contradiction* with oral speech. In section 2.2, I use several examples to demonstrate how *the effects from writing are reproduced in auditory and visual communication*. In section 2.3, I review a case of language change where the written turn actuates the change and directs it, with both the trends identified in previous sections in operation.

It can be argued that some of the observed patterns are somewhat peripheral. It is, however, important that the innovations *exist*, cf. Labov's (2010, 90) claim that "[t]he actuation problem demands that we search for universals in particulars". The mere fact of their existence reveals certain cognitive and social features of the current language situation. To the best of my knowledge, they have received little attention in the literature published so far. In fact, each pattern requires separate and thorough study, so what is presented here is instead a superficial overview, intended to outline a trend and show the possible directions for future, more detailed research.

The aim of this chapter is twofold. First, I intend to produce an overview of some features of the written turn, highlighting its actuality and importance. The focus of sociolinguistics on oral speech is understandable, but a sociolinguistics of writing is also required (cf. Sebba 2009, Lillis 2013). Saussure (2011, 32) called changes triggered by written features "monstrosities" and suggested classifying them as "teratological cases", but in fact change is change, wherever it originates. Second, I intend to propose a partial solution to the actuation problem.

2. Facets of the written turn

2.1. Autonomy of the written speech

Writing has often been considered to be just "a way of recording language" (Bloomfield 1933: 21). According to this view, orthography is a mere representation of phonology, and the features of written texts, for the most part, mirror those of oral speech. Consider in this light the following metalinguistic comment by a Russian Internet user:

> (1)
> *Радуют блондинки, которые пишут "пожáлууууууйста", явно не имея представления об ударениях. *FACEPALM**[41] (bash.org.ru, 2010)
>
> *Raduût blondinki, kotorye pišut "požáluuuuujsta", âvno ne imeâ predstavleniâ ob udareniâh. *FACEPALM**
>
> 'It's fun to observe blondes who write "požáluuuuujsta" ['please'], while clearly having no idea about stress'

Here, the speaker ridicules those users who write "пожáлууууууйста" / "požáluuuuujsta" (classifying them as dumb "blondes"). What is wrong with this spelling? The elongation of a vowel or a consonant is often used in spoken Russian to express emphasis of some kind. The same device is often used in writing, mirroring the emphatic pronunciation. There are, however, certain restrictions on what can be elongated, and since the use of this device in writing is supposed to be a mere recording of language, the restrictions in both registers are defined by the properties of oral speech.

[41] Examples are quoted without any changes, except that stress marks are added where relevant. Deviations from the standard norm are not translated or commented on unless directly relevant to the discussion. The transliteration system used is ISO 9.

As regards vowels, the general rule is that only the stressed vowel can be elongated. The reasons are purely phonetic: if a non-stressed vowel is elongated, while the stressed one is not, the word is likely to sound as if the stress shifts to the elongated vowel, i.e. to the wrong place. Thus, *пожáаалуйста* / *požáaalujsta* is a "licensed" elongation, acceptable in standard written Russian[42]. The general rule has certain exceptions. In some cases, the first vowel can be elongated, even if it is not stressed (*маакóй / taakój* 'such', *бо-ольшóй / bo-ol'šój* 'large'). A pronunciation where the last vowel is elongated (even if it is not stressed), or all vowels are elongated, can also be imagined. They might sound somewhat artificial, but they are not blatantly incorrect[43]. The form *пожáлууууууйста* / *požáluuuuujsta*, however, cannot sound like correct Russian when pronounced. The elongation of *y / u* is unacceptable in this word, both in writing and in speech.

In the unregimented realms of CMC, in blogs, fora, chatrooms and any channels where one can write without being edited by a professional editor, it can occur. Moreover, other instances of the pattern "elongate whatever vowel you want" can be found: consider examples (2–5) (keywords are underlined both in the examples and in the translations).

(2)
<u>убиииивáть</u> таких уродов (vk.com, 2012)

<u>ubiiiiivát'</u> takih urodov

'morons like these have to be <u>killed'</u>

[42] The claims relating to the acceptability of a particular spelling in standard written Russian are based not only on my intuition, but also on verifitcation in the Russian National Corpus (www.ruscorpora.ru, which includes predominantly fiction and non-fiction books), the Integrum database of contemporary Russian press (www.integrumworld.com) and Google books (books.google.com). Details are not provided here for the sake of brevity.

[43] Further exceptions can possibly be found in songs, but, even there, elongation of one and only one non-stressed vowel, if it is neither the first nor the last one is unlikely.

(3)
ой <u>смехооотá</u> какая (pesikot.org, 2007)

oj <u>smehooótá</u> kakaâ

‘oh, how <u>funny</u>’

(4)
<u>лáскоооовый</u>  *такой* (zverushki.tomsk.ru, 2009)

<u>laskoooovyj</u>  *takoj*

‘[it is] so <u>tender</u>’

(5)
матч <u>заеееебúсь</u> (chat at cspl.ru, 2012)

matč <u>zaeeeeebis'</u>

‘the game was <u>fucking cool</u>’

As regards consonants, the restrictions existing in standard language are less clear. Viktorova (2008) provides some limited evidence that the first consonant is elongated more often than other consonants and that sonorants are elongated more often than other types, but the patterns are not as clear-cut as is the case with vowels. Still, CMC obviously shows more variation in the elongation of consonants than traditional channels do, consider:

(6)
и как <u>побббежáла</u> (bash.org.ru, 2012)

i kak <u>pobbbežála</u>

‘and then she <u>really broke into a run</u>’

(7)
оттстáнь (beon.ru, 2011)

otttstán'

‘leave me alone’.

The elongation of a stop consonant such as [b] or [t] is a difficult venture, especially in a [tst] cluster, as in example (8). The phonetic results will resemble stuttering rather than emphasis (although the context in both cases clearly indicates emphasis).

Finally, the Russian alphabet includes two letters that are neither vowels nor consonants, but so-called “signs”, a soft sign ь / ′ and a hard sign ъ / ″. These do not denote any sound and, obviously, cannot be elongated in standard language. In CMC, however, counterexamples can be found, consider:

(9)
брысьььь отсюда (twitter.com, 2012)

brys′′′′ otsûda

‘get out of here’

(10)
ноут вообще вещьььь (bash.org.ru, 2012)

nout voobŝe veŝ′′′

‘a laptop is a cool thing’

(11)
так въъъебать охото ему (clipiki.ru, 2011)

tak v′′′′′′ebat' ohoto emu

‘I want to smack him’

In example (9), the soft sign denotes only the softness of the preceding consonant [s], while in example (10) it has no phonetic value at all and is retained for historical reasons. The hard sign in example (11) does not denote any sound either, indicating instead how the preceding consonant and the following vowel should be pronounced (fulfilling the so-called separating function). Still, speakers find it perfectly possible to express emphasis by elongating these mute letters.

To sum up, the novelty of the observed pattern can be described as follows. Traditionally, in standard written Russian it is possible to express emphasis by mimicking an emphatic device of oral speech, namely, by elongating one or more letters in a word. The letters have to denote sounds that would have been elongated in emphatic pronunciation. In informal CMC, however, it is possible to express emphasis by elongating *any* letters, even if they denote sounds which cannot be elongated in oral speech or do not denote any sound at all.

How could CMC have fostered the emergence of this innovation?

First, through the visual nature of the channel: mostly text, no sound (voice, video and multimodal communication via computers is, of course, possible and is becoming increasingly frequent, but the majority of CMC has been, and still is, text-based, sometimes text-only). This means that traditional prosodic and non-verbal cues are unavailable, and the need to compensate for their absence is strong. In particular, people are eager to convey emphasis by one means or another.

Second, Saussure (2011, 31–32), describing the "tyranny of writing", noted that "[b]y imposing itself upon the masses, spelling influences and modifies language" (31), and that this happens "only in highly literate languages where written texts play an important role". CMC enhances this effect. It has become extremely popular and plays an important role in the daily life of many speakers, in some cases *replacing* auditory communication such as a phone conversation or a face-to-face meeting. Speakers become more accustomed to writing and less dependent on oral speech.

Third, CMC is highly tolerant of deviations from the standard norm, and allows considerable freedom. To provide an example related to the elongation device: in "traditional" written channels (books, press), a letter is seldom repeated more than three times, while in CMC occurrences like

пожалуууууууууууууууууууууууууууууйста
požaluuuuuuuuuuuuuuuuuuuuuuuuuuuuuujsta

are quite possible. There is no editor to suppress this rich variation, and it is this variation which paves the way for innovative patterns to emerge.

Fourth, it is easy to produce an additional letter: just hold a key down for several moments longer (or press it several times). This is easier than writing a letter by hand, although not necessarily easier than elongating a sound in oral speech (it depends on which sound is being elongated).

Why is this innovation important? It shows that people can *write* without thinking of how they *speak*, that the visual channel can become decoupled from the auditory channel. This fact challenges the Bloomfieldian view of writing as a mere "recording" of language (the repercussions of which are still strong among linguists) and supports the view, as expressed by Aaron and Joshi (2006), that writing is as natural as spoken language.

Returning to the actuation perspective, writing as a linguistic domain is becoming more autonomous, and changes that are independent of oral speech are more likely to occur. Contrary to Werry's observation (cited in section 1.4), in this case there is no "straining to make written words simulate speech"; users do not care whether writing simulates speech or not.

2.2. Salience of written speech

In this section, I will discuss a trend which runs counter to the "impulse to produce auditory and visual effects in writing" observed by Werry. While users are certainly

keen to do that, it is also the case that *effects from written speech are reproduced audibly or visually.* Consider example (12):

(12)
Теперь, после достаточного знакомства с интернетом, моя младшая сестра, когда видит или слышит что-то "смешное", то с абсолютно спокойным лицом произносит "хд" и все.. (bash.org.ru, 2010)

Teper', posle dostatočnogo znakomstva s internetom, moâ mladšaâ sestra, kogda vidit ili slyšit čto-to "smešnoe", to s absolûtno spokojnym licom proiznosit "hd" i vse..

'Now that she is pretty familiar with the internet, my younger sister, when she hears something "funny", just keeps an absolutely calm face and says "hd", and that's it..."

The etymology of *хд / hd* is as follows. Both in English and in Russian CMC, the emoticon *xD* (sometimes *XD*) is often used to convey laughter. It is an iconic representation of a laughing face, *D* representing a wide open mouth. In Russian CMC, however, the Cyrillic counterpart *хД* is often used (sometimes lowercase *хд*). It is the result of two different types of transliteration: in Androutsopoulos's (2009) terms, *Д* or *д* is a phonetic transliteration of *D*, while *x* is a visual orthographic transliteration of *x.* Russian *x* denotes sound [h], not [ks], and thus is used due to its visual similarity to English *x*. Importantly, *хД* bears very little resemblance to a laughing face, *хд* even less so. Still, speakers use it, since *xD* requires the additional action of switching the keyboard layout from Russian to English, while *хД* does not. In example (12), the speaker claims that their sister uses *хд* in oral speech. It can be pronounced as [hd], or [hede], or, less likely, [hade] (since the standard name for the letter *x* is [ha]).

It is, of course, possible that the example was invented by the speaker and is not actually true. It was, however, posted on a website (bash.org.ru) where all posts are rated by site visitors, who vote either for (if they find them funny, interesting and plausible) or against them. This one has a rating of more than 25, which means that at least 25 other users did not find the situation described completely unrealistic.

Moreover, other similar statements about the use of [hd] in oral speech can be found on the Internet. Finally, further evidence for the process of "voicing" emoticons can be provided. Consider the following fragment of a rhyme from the same website, bash.org.ru:

(13)
Кот, хуя́се, 0_o, одми́н,
Я наде́юсь не оди́н!

Kot, xuáse, 0_o, odmín,
Â nadéûs' ne odín

'Cat, whatthefuck, 0_o, sysadmin,
I hope I am not alone'

The first line is a list of well-known "old" memes from the website: 0_o is an emoticon conveying surprise (representing the eyes of a surprised person). They are often used together in the form of a similar list at the end of a post (without any connection to the post content). In the preceding lines, the author laments the fact that site visitors are beginning to forget these memes, and then claims he has not, before expresses hope that he is not alone.

Here, it is the metrical foot of the rhyme which is of importance. Most of the previous lines (not cited here) are written in a trochee: DUM-da-DUM-da-DUM-da-DUM. The same clearly applies to the second line of the cited extract: ***Я*** *наде́**е***юсь ***не*** *оди́**и***н!* / ***Â*** *nadéûs' n**e** odí**i**n*. The first line, however, does not fit well: *К**о**т, хуя́**е**, 0_o, одми́**и**н* / *К**o**t, xuá**e**, 0_o, odmí**i**n*, or DUM-da-DUM-da-da-DUM. Both the number of syllables and the structure of the line are wrong. The rhyme can be read correctly only if we assume that the emoticon 0_o functions as a single stressed syllable. In this case the line can be read as *К**о**т, хуя́**е**, **0_o**, одми́**и**н* / *К**o**t, xuá**e**, **0_o**, odmí**i**n*, or DUM-da-DUM-da-DUM-da-DUM.

Most likely, the author reads 0_o as [o], thus ignoring the fact that the first symbol is not *o*, but zero, and that *o* is used here not to denote a sound, but to represent an eye. Alternatively, he may have gauged the suitability of the foot only by looking at the

written lines, almost without trying to read them aloud. If he perceives 0_o as a monosyllabic word, it could act as the required DUM.

Of which novel pattern are these two examples evidence? A linguistic feature which emerged in written speech and is typically associated with writing is borrowed into oral speech. This is not new per se. In Russian, for instance, the set phrase *в кавычках* / *v kavyčkah* has been used for ages (literally 'in quotes', actual meaning 'metaphorically, ironically'). Nowadays, CMC fosters the emergence of new features (in these examples, emoticons *хД* and *0_o*, but similar contexts can be provided for features of other types), which, despite their visual nature, are also borrowed into oral speech.

Why is this important? It shows that the speakers do not perceive emoticons as features confined to the written space, but are ready to invent a way to pronounce them, thus enriching oral speech with devices borrowed from written speech. In example (12), the speaker is aware that something unusual has happened, while in example (13), the speaker uses *0_o* in a rhyme without any explicit comment, being sure that his readers will have no trouble reading it.

2.3. Written speech as an actuator of change

In this section, I will briefly describe a morphosyntactic innovation in Russian which has recently diffused through CMC. For a detailed analysis of the innovation and the evidence for the causal link between its spread and CMC, see Chapter 3.

In the early 2000s, an anti-standard idiom called *Olbanian language* became immensely popular in Russian CMC (see Chapter 6). One of its basic principles is to 'break all rules of orthography', and one manifestation of this principle is the spelling of *o* instead of the unstressed *a* (it violates orthographic but not phonological norms: in unstressed positions both vowels are reduced and sound the same). Thus, nouns with unstressed ending — *дéвушка*, *кни́жка*, *нау́ка*, *мужчи́на* / *dévuška*, *knížka*,

наýка, *мужчи́на* / ‘girl’, ‘book’, ‘science’, ‘man’ — sometimes assume the form *дéвушко*, *кни́жко*, *наýко*, *мужчи́но* / *dévuško*, *knížko*, *naúko*, *mužčíno*.

Nouns ending in *-o* in Russian are strongly linked to the neuter gender. As a consequence of this link, words like *дéвушко* / *dévuško* (originally feminine) or *мужчи́но* / *mužčíno* (originally masculine) can be reclassified as neuter, with their declension and agreement following suit. We are witnessing a rare phenomenon: orthography influencing morphology. The “orthographic neuter” has become frequent and productive (almost losing its connection with Olbanian, the popularity of which is fading). Nouns belonging to other formal classes, not only those ending in unstressed *-a*, are also migrating to the new gender.

The innovation is, of course, used playfully and does not replace any of the three traditional genders (masculine, feminine and neuter), but instead adds a new one. While the original substitution of the letter *o* instead of the unstressed *a* was a purely visual phenomenon, its grammatical consequences affect oral speech as well, cf. standard *тупóй блондинке* / *tupoj blondinke* ‘dumb-DAT.SG.F blonde-DAT.SG’ vs. orthographic-neuter *тупому блондинку* / *tupomu blondinku* ‘dumb-DAT.SG.N blonde-DAT.SG’ or standard *немытый админ* / *nemytyj admin* ‘dirty-NOM.SG.M admin-NOM.SG’ vs. orthographic-neuter *немытое адмúно* / *nemytoe admíno* ‘dirty-NOM.SG.N admin-NOM.SG’.

Moreover, there is anecdotal evidence that speakers sometimes do pronounce non-fully-reduced [o] instead of a reduced [a], imitating the orthographical neuter: *kóšk[o], kís[o], Pólink[o]* instead of standard *kóšk[a]* ‘cat’, *kís[a]* ‘pussycat’, *Polínk[a]*.

As a speaker innovation, the orthographic neuter is not new. It was used in fiction, poetry and the press throughout the 20th century (and the earliest known example dates back to 1888). Two things, however, are important. First, it did not become a community innovation until the early 2000s. Second, prior to that it was used as a salient artistic device, usually adding certain (often pejorative) connotations to the

meaning of a noun (for instance, it could imply the asexuality of a denoted person). Now, on the contrary, having diffused widely enough in CMC, the innovation has lost most of its original semantic connotations, and has become a means of adding some general ironic flavor to one's utterance.

The features of CMC which were instrumental in transforming a rare wordplay into a productive model are as follows. First, CMC provided a written channel of communication: in standard oral speech, *дéвушка* / *dévuška* and *дéвушко* / *dévuško* are indistinguishable. Second, unlike traditional written channels, CMC is both public and unregimented: anybody can write, anybody can read, and nobody edits, i.e. enforces the standard norm. This, as I mentioned in section 2.1, makes the channel tolerant of norm deviations. The public nature of the channel, the fact that it is a multiplex network with a high speed of information exchange, makes the rapid spread of innovations possible.

When the orthographic neuter is finally diffused, both trends identified in the previous sections are in operation. The readiness of speakers to use in writing features that *contradict* oral speech enables the spelling of *o* instead of *a.* The readiness of speakers to *borrow* features from writing into oral speech makes the innovation visible in spoken Russian (since people can pronounce *тупому блондинку* / *tupomu blondinku* instead of *тупóй блондинке* / *tupoj blondinke* or even *kís[o]* instead of *kís[a]*). While the phenomena analyzed in sections 2.1 and 2.2 can be said to be peripheral, the orthographic neuter is frequent, productive, salient and stable.

It is informative to compare the orthographic neuter with a trend observed by Polinsky (2006, 223–224) in what she terms "American Russian", the incompletely learned language of those Russians who moved to USA and became English-dominant in childhood. In American Russian, a shift in the opposite direction happens: neuter nouns are often treated as feminine: *большáя я́блоко / bol'šáâ á́bloko* 'big-NOM.SG.F apple-NOM.SG' instead of standard *большóе я́блоко / bol'šóe á́bloko* 'big-NOM.SG.F apple-NOM.SG'. This shift, however, affects neuter nouns with stem stress (like *я́блоко / á́bloko*) less frequently than nouns with end stress (*лицó* /

licó 'face'). The noun *я́блоко / а́bloko* sounds *а́blok[a]* is thus more easily associated with feminine (for which the ending –a is typical) than *lic[ó]*.

Thus, the situation with morphological gender in Russian resembles what Labov (2010) terms "a fork in the road".[44] Depending on which forces are operating, the linguistic system can follow one route or the other. In this case, the *incomplete acquisition* in American Russian causes nouns to migrate from neuter, the most unstable gender, to feminine, partially destroying the gender system and simplifying it. The *influence of CMC* in mainland Russian causes nouns to migrate from both masculine and feminine to neuter, thus enriching the system with a new "ironic" gender, which complicates the system, and reinforcing the position of the neuter (Zubova 2010).

It has to be noted that orthographic neuter is considerably less salient on a scale of mainland Russian than neuter-to-feminine conversion on a scale of American Russian.

3. Conclusions

Discussing language change, Labov (2010) distinguishes between *triggering events*, which are "particular accidents of history" (Labov 2010: 184), *governing principles*, which constrain the direction of a change, and *driving forces*, which motivate the continuous process of change. Andersen (1989, 8) notes that while linguistic change in some sense is a product of speaker's free will, its development also depends on "the universal principles which govern language use and grammar formation".

In this chapter, I foreground a concept of *selective forces*, related to Labov's notions of governing principles and driving forces and Andersen's of universal principles. A

[44] The difference is that Labov defines "a fork in the road" as an unstable situation when two routes are nearly equiprobable, and the choice depends on some "small forces".

selective force is a factor which either promotes or inhibits the diffusion of a speaker innovation. The main theoretical claim of this chapter is that, in order to explain the actuation of language change in CMC, it is necessary to focus on how CMC changes the selective forces.

This does not presuppose that speaker innovations are of no interest. They are, although in actual fact they seldom become the focus of sociolinguists' attention. Croft (2000, 55) notes that discussions of variation almost never explain how variants arise, and instead most research focuses on their early diffusion. Since in CMC the data are archived and searchable, there are new opportunities to study speaker innovations. The innovations become potentially more searchable, and direct observation of the actuation process, which Milroy and Milroy (1985, 370) claimed to be "difficult, if not impossible", turns out to be possible in some cases.

Still, finding a speaker innovation does not always solve the actuation problem. When emoticons were invented by Fahlman in 1982, they diffused and change happened, but why did it not happen when they were invented in 1967 by Nabokov or in 1887 by Bierce? Why did orthographic neuter in Russian not diffuse in 1888? The list can be continued: variation in language is immense, and one can find a surprising number of speaker innovations at any point in history. For virtually any feature which is presumably new, some older equivalent can be found. Crudely put, at the level of speaker innovations almost everything is possible, but few novel patterns develop into community innovations. Which speaker innovations are successful depends on the combination of actual selective forces.

Moreover, in some cases there may be no single speaker innovation to actuate the change. This is most likely the case with the orthographic neuter: the playful device of converting nouns into neuter has been in use for ages — in the 1960s, 1990s and 2000s. After the spread of Olbanian, however, speakers started using this device with increasing frequency, and, given that selective forces acting against the innovation were now less strong , it diffused rapidly. There was no "zero patient", no single

"author" of the innovation. The idea was obvious enough, and many speakers started using it independently.

The identification and description of selective forces enable us to explain the observed changes, and, potentially, to predict which changes are likely to occur and which are not. Obviously, in order to pursue the ambitious goal of *predicting* changes, sociolinguists also have to achieve a better understanding of the mechanism of speaker innovations (to what extent are they random? can they be predicted?). Likewise, much remains unknown about the mechanism of innovation diffusion through social networks, although this is probably the most widely researched aspect of language change. The mechanism of selection, however, is of most importance, at least if one is interested in the role of CMC, the new communication channel, in linguistic change.

As is obvious from the discussion in section 2, selective forces can be very heterogeneous and include cognitive factors, technological affordances, the structure of social networks and many other things. In this chapter, I offer neither a strict definition of selective force, nor a formalized way of describing it, confining myself to informal descriptions of the written turn.

In section 2.1, I describe how some of CMC's features mitigate a selective pressure from those written innovations which contradict some features of oral speech. This makes written language more independent. In section 2.2, I show how CMC reinforces the pressure to borrow written innovations into oral speech. This makes written language more influential. Finally, in 2.3 I show how CMC removes selective pressures from the orthographic neuter, and how the two trends outlined in 2.1 and 2.2 play a role in its successful diffusion.

The partial solution to the actuation problem, promised in section 1.4, can be formulated as follows. *A particular change X can take place in a particular language at a given time because selective forces, which operate in a given language, in a given society and at a given time, and which play an important role in determining*

which speaker innovations will diffuse and which will not, promote particular speaker innovations, namely, those that can trigger the change X.

One factor which affects how selective forces operate is the communication channel. As this chapter shows, some of the recent innovations in Russian were successfully actuated because of the availability of CMC.

Acknowledgments

I am grateful to all the participants of the "F4" workshop (Bergen, 2012) who participated in the discussion of this paper, in particular to Sali Tagliamonte. This work also benefited from comments by Alexander Piperski and Arkadij Avdohin.

References

Aaron, P. and R. Malatesha Joshi. 2006. "Written language is as natural as spoken language: a biolinguistic perspective." *Reading Psychology* 27: 263–311.

Andersen, Henning. 1989. "Understanding linguistic innovations". In *Language change: contributions to the study of its causes*, edited by Leiv Egil Breivik and Ernst Håkon Jahr, 5–28. Berlin: Mouton de Gruyter.

Andersen, Henning, ed. 2001. *Actualization: linguistic change in progress.* Amsterdam, Philadelphia: Benjamins.

Androutsopoulos, Jannis. 2011a. "Language change and digital media: A review of conceptions and evidence". In *Standard languages and language standards in a changing Europe*, edited by Kristiansen Tore and Nikolas Coupland, 145–161. Oslo: Novus.

Androutsopoulos, Jannis. 2011b. "From variation to heteroglossia in the study of computer-mediated discourse." In *Digital Discourse: Language in the New Media* edited by Crispin Thurlow and Kristine Mroczek, 277–298. Oxford: OUP.

Baron, Naomi. 1984. “Computer mediated communication as a force in language change.” *Visible language* 18 (2): 118–141.

Baron, Naomi. 2005. “The written turn.” Review of *The English writing system*, by Vivian Cook. *English Language and Linguistics* 9:359-376

Bloomfield, Leonard. 1933. *Language* (revised from 1914 edition). New York: Holt.

Cheshire, Jenny, David Agder and Sue Fox. 2012. “Relative *who* and the actuation problem.” *Lingua*, Available online 9 January 2013, doi 10.1016/j.lingua.2012.11.014.

Collot, Milena and Nancy Belmore. 1996. “Electronic language: A new variety of English”. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 13–28. Amsterdam: Benjamins.

Croft, William. 2000. *Explaining language change: an evolutionary approach.* Harlow: Longman.

Dresner, Eli. 2005. “The topology of auditory and visual perception, linguistic communication, and interactive written discourse.” *Language@Internet* 2, article 2. http://www.languageatinternet.org/articles/2005/161, accessed January 23, 2013.

Fahlman, Scott. ?2002. “Smiley lore.” http://www.cs.cmu.edu/smiley/history.html, accessed 20 January 2013.

Ferrara, Kathleen, Hans Brunner and Greg Whittemore. 1991. “Interactive written discourse as an emergent register.” *Written Communication* 8 (1): 8–34.

Herring, Susan. 2012. “Grammar and electronic communication.” In *The Encyclopedia of Applied Linguistics*, edited by Carol Chapelle. Hoboken, New Jersey: Wiley-Blackwell.
http://onlinelibrary.wiley.com/book/10.1002/9781405198431

Herring, Susan. 2003. “Media and language change: introduction.” *Journal of Historical Pragmatics* 4 (1): 1–17.

Jones, Graham and Schieffelin, Bambi. 2009. “Enquoting voices, accomplishing talk: Uses of be + like in instant messaging.” *Language & Communication* 29: 77–113.

Ko, Kwang-Kyu. 1996. “Structural Characteristics of Computer-Mediated Language: A Comparative Analysis of InterChange Discourse.” *Electronic Journal of Communication* 6 (3)

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

Labov, William. 1994. *Principles of linguistic change, Internal factors*. Hoboken: John Wiley & sons.

Labov, William. 2010. *Principles of linguistic change, cognitive and cultural factors*. Hoboken: John Wiley & sons.

Lillis, Theresa. 2013. *The sociolinguistics of writing.* Edinburgh: Edinburgh University Press.

Maslova, Elena. 2008. “Unidirectionality of grammaticalization in an evolutionary perspective.” In *Studies on grammaticalization*, edited by Elisabeth Verhoeven, Stavros Skopeteas, Yong-Min Shin, Yoko Nishina, Johannes Helmbrecht, 15–24. Berlin, New York: Mouton de Gruyter.

Milroy, James, and Lesley Milroy. 1985. “Linguistic change, social network and speaker innovation.” *Journal of Linguistics* 21 (2): 339–384.

Paolillo, John. 2001. “Language variation on Internet Relay Chat: A social network approach.” *Journal of Sociolinguistics* 5 (2): 180–213.

Polinsky, Maria. 2006. “Incomplete acquisition: American Russian.” *Journal of Slavic Linguistics* 14 (2): 191–262.

Saussure, Ferdinand. 2011. *Course in general linguistics*. Translated by Wade Baskin, edited by Perry Meisel and Haun Saussy. New York: Columbia University Press.

Sebba, Mark. 2009. Sociolinguistic approaches to writing systems research. *Writing Systems Research* 1 (1): 35–49.

Tagliamonte, Sali. 2006. *Analysing sociolinguistic variation*. Cambridge: CUP.

Tagliamonte, Sali, and Derek Denis. 2008. "Linguistic ruin? LOL! Instant messaging and teen language." *American Speech* 83(1): 3–34.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: OUP

Viktorova, M. 2008. "Èmotivnyj aspekt prodleniâ soglasnyh zvukov (na materiale pesen V. Vysockogo i Ž. Brelâ." [The affective aspect of elongating the consonant sounds (on the data from the songs of Vladimir Vysotsky and Jacques Brel] *Izvestiâ Rossijskogo gosudarstvennogo pedagogičeskogo universiteta imeni A. I.* Gercena 60: 68–73.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. *Empirical foundations for a theory of language change.* Austin: University of Texas Press.

Werry, Christopher. 1996. "Linguistic and interactional features of Internet Relay Chat. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 47–61. Amsterdam: Benjamins.

Wolfram, Walt. 1991. "The linguistic variable: fact and fantasy." *American speech* 66 (1): 22–32.

Yates, Simeon. 1996. "Oral and written linguistic aspects of computer conferencing: A corpus based study" In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 29–46. Amsterdam: Benjamins.

Zimmer, Ben. 2007. “The prehistory of emoticons.” *The Language Log*, September 21, 2007. http://itre.cis.upenn.edu/~myl/languagelog/archives/004935.html

Zitzen, Michaela, and Dieter Stein. 2004. “Chat and conversation: a case of transmedial stability?” *Linguistics* 42(5): 983–1021.

Zubova, Lûdmila. 2010. “Ironičeskaâ grammatika: srednij rod v igrovoj neologii.” [Ironic grammar: neuter gender in playful neology] *Voprosy âzykoznaniâ* 6: 16–25.

V

5. “You were typing something, I interrupted”: Play with conversation structure in quasi-synchronous computer-mediated communication

1. Introduction

1.1. Research question

In this article, I address the following research question: do the properties of computer-mediated communication (CMC) affect how its users play with language? Do any novel patterns emerge, are any novel devices being created? These questions can be viewed as part of a larger research question — how does CMC affect the linguistic behavior of its users?

I will be focusing on one language (Russian), one type of language play (play with conversation structure) and one type of text-based CMC, the so-called “quasi-synchronous computer-mediated communication” (QS-CMC). In the following sections, I explain why play with conversation structure is worth investigating and what QS-CMC is.

1.2. Language play and CMC

For clarity’s sake, several definitions need to be introduced. *Language play*, or *linguistic humor*, is humor related to the surface structure form. Those formal linguistic elements which are necessary in order to create the comic effect cannot be changed (i.e. in retelling or translation) without destroying the humor. Another important notion is *humor in interaction* (Norrick 2010), or *conversational humor*, defined by Dynel (2009, 1286) as:

> various verbal chunks created spontaneously or repeated verbatim for the sake of amusing the recipient, either directly contributing to the semantic content of the ongoing conversation or diverting its flow into a humorous

> mode/ frame /key, in which speakers need not genuinely mean what their humorous verbalisations convey.

Dynel (2009) proposes a classification of most common types of conversational humor that I rely upon when using terms like *teasing*, *putdown*, *butt of a joke* etc. Both language play and conversational humor can be viewed as subtypes of *verbal humor*, any humor conveyed by language rather than by non-verbal stimuli (Dynel 2011a). Obviously, language play and conversational humor cannot be viewed as taxa of a single classification, so an instance of ludic behavior can fall under either one of these categories, or none, or both at once.

My focus is on humorous actions of the latter type, those that can be considered *both* conversational humor and language play. In other words, I investigate how speakers play with the conventions of conversational organization[45] in a quasi-synchronous communication channel.

In the early days of CMC studies, Baym (1995) claimed that computer-mediated communication (CMC) "is rarely seen as a means of making people laugh", and that little work up to that time had examined the role of humor in CMC.

Today, the former claim no longer seems valid. Numerous studies acknowledge that humor in general and language play in particular are welcome in various CMC channels. To give a few examples: Werry (1996) mentions this re IRC, Simpson (2005) re synchronous chats in general, Anis (2007) and Plester and Wood (2009) re texting, and Danet (2001) and Crystal (2006) re CMC in general.

Nor is Baym's latter claim as true as it once was, as quite a few studies have investigated various aspects of humor in CMC: play with identity (Bechar-Israeli 1995); the role of humor in task-solving (Morkes, Kernal and Nass 1998), in

[45] Due to the vagueness of the linguistic humor/language play concept, it can be argued that at least some of my examples do not fall under this category. This is largely a matter of definition and does not affect the substance of the argument.

provoking and mediating conflict (Rouzie 2001); or numerous aspects all at once (Danet, Ruedenberg-Wright and Rosenbaum-Tamari 1997, Danet 2001).

Most of this research, however, examines the role of *humor in CMC*, not the role of *CMC in humor*. In other words, it does not empirically address questions such as the one outlined at the beginning of this article: does CMC change patterns of playful linguistic behavior in conversation?

Studies of CMC which are more language-oriented often pay some attention to language play (Yates and Orlikowski 1993; Werry, 1996; Herring, 1999; Hård af Segerstad 2002; Simpson 2005; Mechkovskaia, 2006; Crystal, 2006), but it is seldom a primary focus of research. Exceptions include studies of the creative use of orthography (Su 2003; Gottlieb 2010), second-language acquisition studies (Belz and Reinhardt 2004, Warner 2004, Vandergriff and Fuchs 2009; the latter also empirically challenges the common assumption that CMC promotes language play) and an exploration of conversational humor in a specific CMC genre of online sports commentary (Chovanec 2011).

In some cases, researchers of humor explicitly rely on CMC as a data source (Thielemann 2011, Popescu 2011, Dynel 2011b). Taking the data from the Internet, however, does not per se problematize the role of the Internet in shaping the linguistic structure of humor.

Why is it important to study how CMC affects language play? First, play is important per se, it is a ubiquitous and prominent aspect of linguistic behavior. If its patterns undergo any changes when speakers start using a new communication channel, the reasons for these changes and their nature are of great interest to linguists.

Second, as observed by Sannikov (2002), language play can be used to study the properties of the language itself. Sannikov argues that language play is a linguistic experiment performed by language speakers. Since there is always some kind of anomaly underlying language play, we can use the data from this ongoing ubiquitous experiment in *anomaly* to study what is *normal.* Thus, humorous distortions of

conversation structure might improve our understanding of what conversational norms exist in QS-CMC.

1.3. Quasi-synchronous CMC

The overwhelming majority of instant messengers, chats and similar communication programs can be classified as quasi-synchronous, since, as Garcia and Jacobs (1999, 339) note, "although posted messages are available synchronously to participants, the message production process is available only to the person composing the message", and, consequently, message production is not synchronous with message transmission.

Herring (2007), however, in outlining a classification scheme for computer-mediated discourse, proposes another conceptualization. Speaking of medium factors, she singles out *synchronicity* and *transmission* (one-way vs. two-way) as separate factors. Synchronicity is defined as the requirement that sender and addressee(s) are logged on simultaneously. Transmission can occur either message-by-message, as described above by Garcia and Jacobs and as happens in most instant messengers, or keystroke-by-keystroke, as in Google Wave or the VAX Phone system (Anderson, Beard and Walther 2010). Following Cherny (1999), Herring terms the former *one-way transmission* and the latter *two-way transmission*.

This classification, however, suffers from several drawbacks. First, not all instant messengers require both sender and addressee(s) to be logged on. True, it is usually assumed that they are, but in ICQ, for instance, it is perfectly possible for Alice to send a message to Bob while he is not logged on. When Bob logs on, he receives the message, even if Alice has already logged off. Bob can then, in turn, send a reply to Alice, even though she is absent. ICQ, then, does not satisfy Herring's formal criterion of synchronicity, but it hardly makes sense, for instance, to classify ICQ as asynchronous, and Skype (where Alice and Bob do indeed have to be logged on simultaneously for a message to arrive) as synchronous.

Second, the labels “one-way” and “two-way” are opaque. They lead one to think of unidirectional (e.g. broadcaster—>audience) vs. bidirectional communication rather than of message-by-message vs. keystroke-by-keystroke transmission. Third, important parameters such as speed of transmission and program interface are completely left out.

In my view, what we intuitively understand as synchronicity is, in fact, a complex feature which depends on several parameters and which constitutes a cline. I operationalize this view by the following working definition: the synchronicity of (computer-)mediated communication depends on:

— how quickly the messages are delivered from the sender to the addressee;

— whether the interface facilitates immediate reading of the delivered messages and immediate responses to them;

— to what extent the transmission is synchronized with production and with perception (for text-based CMC — whether it occurs message-by-message or keystroke-by-keystroke);

— whether the interlocutors are required or expected to be simultaneously “present”.

The point on this cline where e-mail lies, for instance, can be labeled as *asynchronous*: message delivery is fast, but not necessarily instantaneous; interface does not facilitate immediate reply: typically, one has to click “Reply”, open a new window, type the message, and then send the message; no co-presence is expected. ICQ, IRC, Skype in text mode etc. can be considered *quasi-synchronous:* messages are delivered immediately; to respond it is enough to type the message in the same window and press Enter; no keystroke-by-keystroke transmission; co-presence is expected (see also Zitzen and Stein 2004 for the detailed discussion of the properties of quasi-synchronicity). Software like the VAX Phone system can be called *synchronous*.

Note, however, that communication can be more asynchronous than e-mail (e.g. snail mail) or more synchronous than synchronous chats (e.g. video communication, where message transmission occurs in real time and is enriched by non-verbal cues which accelerate perception and facilitate the organization of turn-taking).

As I mentioned above, this article deals with the point of the synchronicity cline labeled "quasi-synchronous". Previous research has shown that the conversation structure in QS-CMC has its peculiarities (Baron 2010, Berglund 2009, Simpson 2005, Zitzen and Stein 2004, Garcia and Jacobs 1999, Herring 1999; Burkhaller, Smith and Cadiz 2000). Here, I address the role of QS-CMC in shaping the playful conversational behavior of its users.

1.4. Constraints and affordances

When describing how users adapt to the properties of a mediated communication, it is often convenient to use the notions of affordances (or similar, but differently labeled notions), cf. Hutchby 2001, Hård af Segerstad 2002, Lee 2007. Affordance is a certain opportunity for action in a given environment, as perceived by a given user. Likewise, it is often convenient to talk also of constraint, the impossibility of performing a certain action. The choice of label (constraint vs. affordance) is rather arbitrary and depends on one's point of view (we can say that in QS-CMC a sender can use the affordance of not showing their message to their interlocutor until it is finished; but it is equally reasonable to say that the addressee faces the constraint of not being able to see the message-in-progress until it is sent).

In Section 3, I demonstrate the influence of QS-CMC on language play structure by singling out specific affordances offered by this environment and analyzing their use in play with conversation. I use the word *affordance* as a generic term for both affordances and constraints. Obviously, quasi-synchronicity is the principal affordance that manifests itself in one way or another in all the examples, most

straightforwardly in examples 1–9. Other related or unrelated affordances, however, are singled out and discussed wherever possible.

In section 2, I outline the methodology and the data source, and in section 4 I offer some conclusions.

2. Materials and methods

2.1. Theoretical perspective

To analyze conversational structure, I employ methods from conversation analysis (CA). It is to this approach that we owe the attention to the turn-taking mechanism and related aspects of interactive discourse organization: the focus on log data; the tradition of describing the conversation in terms of the participants' own intentions and understandings; and the data-driven methodology (analyze the interaction; find a pattern; explain it).

CA has been productively applied to mediated communication, starting from phone conversations (see Hutchby 2001 for an overview) and proceeding to CMC (see González-Lloret 2011 for a brief review). From the very beginning, CA was also envisaged as an approach that was suitable for the analysis of playful conversations (see, for instance, Sacks 1974). In keeping with its methodological spirit, I do not engage with any major theory of (verbal) humor. My approach to humor can be called methodologically impure[46]: I provide intuitive explanations of what constitutes comic effect, sometimes using concepts garnered unsystematically from various theories. Such, however, is the nature of my data: different affordances stimulate very different types of language play, and it is not surprising that models from different frameworks have to be used to understand them better.

[46] The same can be said of CA itself, but it works (ten Have 1986).

The main difference between this study and more traditional implementations of CA is the nature of the data. The inherent drawback in my examples is that they are all extremely decontextualized (see next section). However, even the short conversation extracts that are available make it possible to offer some explanations for the patterns observed.

2.2. Data source

All examples come from the website "Runet Quote Database" http://bash.im (previously http://bash.org.ru and more widely known under this name), a Russian version of the "Quote Database" http://bash.org. The website is a collection of quotes extracted from real conversations, with a large share of the quotes originating from QS-CMC conversations. Users submit what they regard as funny extracts from their conversations to a section on the website called "Abyss". The submitted quotes are rated by site visitors, who vote either for or against them. The quotes are reviewed by editors, who approve (and sometimes edit) the quotes they like — these appear on the main page. All quotes that earn a user score of at least 30 (until recently — of at least 25) are stored in the section "The best of Abyss".

The website is one of the most popular entertainment websites on the Russian Internet, with about 300 000 visitors daily.[47] As of January 6 2013, 38 513 quotes were approved, 86 were pending and more than 240 000 quotes were archived in "The best of Abyss".[48] To put this in perspective, the original English website had 20 945 quotes approved and 386 quotes pending (there is no section similar to "The best of Abyss").

[47] http://www.liveinternet.ru/stat/bash.org.ru/index.html?period=month, accessed 6 January 2013.

[48] The actual number of quotes available on the website is smaller, since they are deleted after one year's storage.

The "Runet Quote Database" is not the most typical data source, but it is quite useful for investigations of humor in CMC, since all the preliminary work of gathering a huge amount of data (otherwise poorly accessible) has already been done.

One drawback of using it is that we cannot know for sure whether the data are real: some quotes could be fake. For my study, I use quotes from "The best of Abyss" section. First, the editors claim not to edit anything there, and second, all the quotes there have a user score of at least 25. Since site visitors generally dislike "non-genuine" quotes and vote against them,[49] it might be stated that each quote in my selection was considered genuine by at least 25+1 independent reviewers. Even if some quotes are not actually real, they look like they could have occurred in real conversation, and thus are almost equally informative for my purposes.

Another drawback is that examples are decontextualized. In most cases, nothing is known about who the participants are (which is important from an ethnographic point of view), what software and hardware they are using (which is important for a better understanding of affordances), what the broader context of their communication is, who found the conversation funny and posted it on bash.im (one of the participants or a third person) etc. In most cases, some of this information can be reconstructed; when analyzing the conversation, I rely on the most plausible reconstruction.

The lack-of-context drawback is partially compensated by lack of the observer's paradox. The examples presumably originate from natural conversations, where speakers were not thinking that their utterances will subsequently be analyzed by a researcher.

Despite the name of the website and the appeal to users to submit extracts from real conversations, some "quotes" are not quotes stricto sensu, but rather witticisms, humorous observations, comments on other quotes or even appeals to other site

[49] This is known from the appeals "stop posting fakes! " that are sometimes posted on the website, and, in their turn, receive high user scores, as well as from discussions on various fora.

visitors. Such quotes can also secure a very positive rating, and, since they tend to be quite informative, I did not exclude them from the analysis (see examples 5 and 6).

The examples were collected during monitoring of the website from September 2009 to December 2012 (no less than 50000 quotes were reviewed, exact estimate is not possible for technical reasons).

I provide almost no numbers (for instance, how frequent the observed playful devices are), since my data do not allow for any reasonable quantification. First, the original sample is not a *random* collection of texts. It was filtered first by posters, who decide what to post, and then by voters. It is possible, for instance, that device A occurs two times and device B one time, since an extract containing another instance of B was too similar to the one already posted in "Abyss" and thus received low score.[50] In other words, bash.im provides a very good data source if one asks which patterns *exist* in Russian CMC. It is, however, not quite clear how representative the website can be if one asks how *frequent* the patterns are. Addressing such questions, great care has to be taken and special methods probably have to be devised. I did not do this, since the features I am interested in, playful devices, are poorly quantifiable per se. It is often difficult to determine what should be considered various instances of the same device and what should not. Moreover, it is beyond the scope of this article. Here, I intend to show the *existence* of certain novel and CMC-specific patterns of play with conversation structure, and for demonstrating existence even a single example can be enough. As Androutsopoulos (2011b, 280) put it, it is important not to "shy away from the importance of singular, unrepeated instances of linguistic difference if and as used in a strategic, yet non-quantifiable way".

[50] I am grateful to Martin Paulsen for pointing out this observation to me.

2.3. Examples and translations

Examples are provided as they were posted on bash.im. Any additional information (timestamps, info messages) that was available has been kept. The original nicknames in the source data are often anonymized, usually to *xxx* and *yyy*, sometimes replaced with *he*, *she* or *I*.

For all the examples, the original Russian version, a transliterated version (I use ISO9, the only major Romanization system for Cyrillic that is fully reversible) and an English translation are provided. Glosses are not provided in the interests of brevity and simplicity, and also since they are not necessary to understanding the conversation structure.

Most of the deviations from the standard norm in the Russian examples are not reproduced in translations, since they are not relevant to the discussion.

Discussing the examples, I often refer to the speakers as he or she, although their gender is not explicitly mentioned. That means that the gender of the speaker is known either from a personal name or a morphological gender marker. When gender is unknown, I use “they” to refer to a speaker.

3. Affordances of QS-CMC and their use in play

3.1. Info message

One relevant affordance, originally created by software designers to compensate for the *quasi*-synchronicity of communications in chatrooms and instant messengers and to make those communications more real-time, is the *info message*. Info messages are automatically generated by software and inform a participant (or several participants at once) that somebody is typing them a message (obviously, other types of software-generated messages exist, but in this text an “info message” is hereafter understood only in this narrow sense). While the person receiving an info message may know

that they are (potential) addressees, the message-in-progress itself remains unavailable to them. This opens up scope for various types of playful interaction: consider, for instance, example (1).

(1) (2009)

1	Ксюша печатает Вам сообщение (20:34)
2	Ксюша печатает Вам сообщение (20:46)
3	Ксюша печатает Вам сообщение (20:59)
4	Я: Ты там мне что, поэму пишешь?
5	Ксюша печатает Вам сообщение (21:06)
6	Ксюша печатает Вам сообщение (21:17)
7	Я: Чувствую, ща твоего бреда с ошибками начитаюсь...
8	Ксюша печатает Вам сообщение (21:28)
9	Ксюша: Привет, Даня!

1	Ksûša pečataet Vam soobŝenie (20:34)
2	Ksûša pečataet Vam soobŝenie (20:46)
3	Ksûša pečataet Vam soobŝenie (20:59)
4	I: Ty tam mne čto, poèmu pišeš′?
5	Ksûša pečataet Vam soobŝenie (21:06)
6	Ksûša pečataet Vam soobŝenie (21:17)
7	I: Čuvstvuû, ŝa tvoego breda s ošibkami načitaûs′...
8	Ksûša pečataet Vam soobŝenie (21:28)
9	Ksûša: Privet, Danâ!

1	Ksûša is typing you a message (20:34)
2	Ksûša is typing you a message (20:46)
3	Ksûša is typing you a message (20:59)
4	I: What are you up to there, are you writing me a [long] poem?
5	Ksûša is typing you a message (21:06)
6	Ksûša is typing you a message (21:17)
7	I: I sense I’ll be reading lots of your ravings with [orthographical] errors soon...
8	Ksûša is typing you a message (21:28)
9	Ksûša: Hi Danâ!

Over a period of 25 minutes, Danâ receives three info messages (1–3) telling him that Ksûša is typing him a message. Devoting such a long time to a single message does not fit with standard practice in QS-CMC, where users usually opt for short messages that do not require time-consuming editing. Danâ's sarcastic response (4) suggests that he interprets Ksûša's unusual behavior as follows: the message is extremely long, and probably also of high formal complexity (i.e. poetic). His next response (7), however, shows that he has abandoned the latter hypothesis (or, more likely, never considered it seriously). Still, he is obviously expecting to receive a long text. This is entirely natural, especially given that the info messages remain on his screen, amplifying the impression that a long message is being composed. When the message finally arrives (9), it turns out to contain nothing but a short common greeting, which confounds the expectations of both Danâ and the external reader (speaking in terms of incongruity-based humor theories, it resolves an incongruity created by preceding exchanges) and thus creates a comic effect.

Various explanations of what Ksûša was actually doing are possible. The most plausible account, in my view, looks as follows: on multiple occasions (1, 2, 3, 5, 6, 8) she began typing a message to Danâ, but every time she stopped for some reason, erasing the message-in-progress, before finally sending it (9)[51]. Aborting one's own message production is not uncommon in QS-CMC, and there are many possible reasons: a sudden increase in the priority of another task; the desire to postpone the contact, for whatever reason; choice of words etc. Ksûša's behavior, however, looks unusual even by QS-CMC standards: she aborts message production at least five times within an hour, she ignores Danâ's comments and questions elicited by her actions, and her final message contains no excuse or explanation. The comic effect persists even if we accept that Ksûša did not actually spend 54 minutes writing this message (as the timestamps suggest), since five failed attempts are, in any case, too many for a short greeting.

[51] Alternatively, she might be experiencing problems with her network.

The conversation and the comic effect are made possible by the quasi-synchronicity and the availability of info messages. These two factors also result in the emergence of some other interesting effects. First, note how the conversation opens. There is no greeting in Danâ's first turn (4), although Ksûša does greet him in (9), thus implying that they have probably not spoken for some time and a greeting is in order. More interestingly, there is no intentional summons for Danâ to speak, such as a welcoming gesture or the ringing of a telephone: he chooses to respond to the messages that are generated by a machine, presumably independently of Ksûša's will. (In principle, a description which assumes that Ksûša intentionally started and finished working on messages-in-progress to attract Danâ's attention is not impossible. Alternatively, it is also possible that she was not aware that her interlocutor was seeing these messages, since their presence usually depends on the settings of the client). Danâ's behavior justifies CA's refusal to focus exclusively on the speaker, treating the hearer "as a figment of the speaker's imagination" (Goodwin and Heritage 1990, 292). In this example, Danâ, discovering that he is an addressee, or rather is going to become one, manifests his impatience to see the utterance-in-progress. This desire also fits with the basic principle of turn-taking: no gaps, no overlaps. In this case, Ksûša's decisions not to send her messages create a gap (which only becomes visible due to the availability of the info messages), and that is not easily tolerated by Danâ.

Example (2) also starts with a potential addressee responding to an info message.

(2) (2009)

1	yyy Печатает.
2	xxx: угу
3	yyy: Ванька! ты тута?
4	xxx: нормально)
5	yyy: круто! как твои делишки?
6	xxx: ладно, давай я буду отвечать после твоих реплик?))
7	yyy: ...

1	yyy Pečataet.
2	xxx: ugu
3	yyy: Van′ka! ty tuta?
4	xxx: normal′no)
5	yyy: kruto! kak tvoi deliški?
6	xxx: ladno, davaj â budu otvečat′ posle tvoih replik?))
7	yyy: ...

1	yyy Is typing.
2	xxx: yeah
3	yyy: Van′ka! are you here?
4	xxx: all’s well :)
5	yyy: cool! how are you doing?
6	xxx: Right, now I’ll start replying after your turns, ok? :))
7	yyy: ...

Seeing the info message (1), xxx responds with a *yeah* (2), apparently unrelated to any question or statement. It finds its first pair part, however, when yyy proceeds with the question *are you here?* (3), and it becomes clear that line 2 can be viewed as containing an answer that anticipated its question. Lines 4 and 5 confirm that xxx is intentionally and successfully exploiting the playful device of answering a question before it is asked.

As a result, the conversation consists of *inverted* adjacency pairs: second pair parts (2, 4) precede the first ones (3, 5).

Again, the play becomes possible due to quasi-synchronicity. It seems likely that xxx is managing to type and send his messages (2, 4) while yyy is busy composing theirs and not paying full attention to the screen. However, even if yyy does pay attention, notices and reads xxx’s messages, they fail to recognize them as potential answers to the upcoming questions. This is not so surprising, given that *yeah* and *all’s well* (literally ‘normally’, a common answer in Russian) can serve many functions in a conversation and appear (at first!) without any context.

Info messages also play an important role here, since xxx uses the first one to open the conversation (cf. example 1). Moreover, he probably continues to rely on them to

time the sending of his messages (2, 4), although the person who posted the extract to bash.org.ru must have deleted them (which is understandable: a simple text-based extract would have failed to convey all the intricacies of the timing of messages anyway).

The meaning of xxx's replies and their ludic nature are revealed in full after his explicit promise to stop anticipating questions (6). It is most probably not random that xxx chooses to stop his performance at this point. While the adjacency pairs in the opening of a conversation are highly formulaic and thus rather easily predictable, this is not the case with the conversation itself. Parts of conversation that involve the actual transmission of information are much less predictable than the phatic parts.

A somewhat similar playful device is used in example (10). While there are no inverted adjacency pairs, there are "reconstructed" ones: unable, for some reason (troubles with encoding or keyboard layout), to read questions, speakers manage to guess them and provide appropriate second turns. Note that in (10) the speaker also achieves full success only with the first two questions.

Interestingly, a play of this "I-know-what-you-will say" genre can also fail, consider example (3).

(3) (2010)

1	*yyy печатает*
2	xxx: Нормально
3	yyy: не угадал

1	*yyy pečataet*
2	xxx: Normal′no
3	yyy: ne ugadal

1	*yyy is typing*
2	xxx: All's well
3	yyy: You guessed wrong

Conversation begins in the same way as in the previous example. Seeing info message (1) xxx attempts to predict yyy's turn and answer with an appropriate second pair part in advance (2). The expected first pair part does not follow, however, since yyy immediately recognizes xxx's intention and takes part in the play, claiming that xxx has failed to predict the contents of the message-in-progress (3). This claim reverses the direction of the teasing, and, unlike example (2), it is xxx who becomes the butt of the joke.

It is possible that the prompt reaction of yyy is facilitated by their familiarity with humor of this kind. For instance, yyy might have participated in such interactions themselves or might have seen example 2 (or other similar extracts) at bash.im. The latter is not unlikely, given the immense popularity of the website and the fact that either yyy or xxx also use it.

Speakers are not unaware of their use of info messages, as is evidenced by the one-line example (4).

(4) (2010)

xxx: ты что-то печатала, я тебя перебил.
xxx: ty čto-to pečatala, â tebâ perebil
xxx: you were typing something, I interrupted

Apparently, xxx has sent a message to his interlocutor (let us dub her yyy) while she was typing a message to him (and this fact could have become known to him only through an info message). As follows from xxx's apology, it must be the case that yyy did not finish the message-in-progress; moreover, she must have stopped typing it.

The situation is similar to an *interruption* in an oral face-to-face (FtF) conversation, and xxx makes use of this similarity, offering a mock apology. The comic effect is created for the most part by the fact that the interruption was not real. The sequence of actions of the interlocutors is the same in a typical FtF interaction and in this example: Alice starts producing an utterance, Bob produces an utterance while Alice's is yet to be finished, and this causes Alice to stop producing her utterance. The effect of Bob's interrupting utterance in oral speech, however, is different from the one it can have in QS-CMC.

In synchronous oral speech, it is difficult for two people to produce messages simultaneously due to the physical properties of the communication channel, but in QS-CMC it is perfectly possible. The "interruption" did not make it impossible for yyy to continue message production.

On the other hand, the "interruption" did indeed cause her to stop typing her message. Most probably, she stopped to read it, and then later did not resume typing, deciding either not to finish the message, or to change it, taking into account what xxx had just said, or for some other reason. Cf. Beißwenger's (2008, 14–15) analysis of video recording of the two chat participants modifying their work on the message-in-progress (continuing it, or stopping typing, or deleting the text) with respect to the interlocutor's behavior (typing, or being silent, or sending a message which is presumably not a complete utterance etc.).

It has been claimed that simultaneous talk, such as overlaps and interruptions, is impossible in QS-CMC (Garcia and Jacobs 1999, 346). As regards overlaps, the truth value of this statement depends on how *overlap* is understood. Simultaneous production of utterances is possible, see examples (1)–(4). This was, for instance, convincingly shown by Beißwenger (2008) and by Smith and Gorsuch (2004), who studied not only the text logs of chats, but also real-time videos that captured everything that had happened on the screens of participants. However, simultaneous production of utterances that are immediately available to all speakers (as in oral speech) is indeed impossible. As regards interruptions, while physically splitting the

utterance-in-progress with another utterance is, of course, impossible, interruption-like situations, as in example (4), do emerge in QS-CMC.

González-Lloret (2011, 312) states that "there is no competition for the right to submit a message" in QS-CMC, but xxx in examples (4) behaves as if there is such a competition, and he inadvertently deprives yyy of her right to submit a message. We do not know how serious he is. He may just be joking, or he may be completely serious (in this case it must have been yyy who viewed the utterance as funny and posted it on bash.im). Most likely, however, he is half-serious: having observed that there exists a pattern in CMC that is similar to an oral-speech pattern, but not identical, he, in his desire to amuse his interlocutor, behaves as if it were entirely identical.

The examples (1)–(4) contradict Zitzen and Stein's (2004, 993) observation that in chat, "there are no possibilities for monitoring the temporary suspension of the right to speak". Info messages provide such a possibility, and speakers use it (although sometimes in unexpected ways), thus challenging another claim of by Zitzen and Stein, namely, that overlaps are the norm in composing messages. While it is undoubtedly true that overlaps in producing messages often occur, example (4) shows that they are not always viewed as norm, and sometimes users try to avoid them.

Metalinguistic comments in examples (5) and (6) are also related to the "no gaps, no overlaps" principle.

(5) (2009)

Умение слушать трансформировалось в умение ждать, пока в аське горит "Печатает ..."

Umenie slušat′ transformirovalos′ v umenie ždat′, poka v as′ke gorit "Pečataet ..."

The ability to listen transformed into the ability to wait while ICQ is showing the "Is typing..." message

(6) (2009)

БЛя.... у меня на 4м году администа вырабатался комплекс... когда мне долго пишут... всмысле аська отображает что мне пишут, так вот когда мне пишут долго сотрудники, у меня терпения не хватает и я обычно просто заглядую к ним через РАдмин... а вот когда мне долго пишут те у кого я не являюсь админом и я немогу посмотреть что там делается... сцуко... я нервничаю...))

BLâ.... u menâ na 4m godu adminsta vyrabatalsâ kompleks... kogda mne dolgo pišut... vsmysle as′ka otobražaet čto mne pišut, tak vot kogda mne pišut dolgo sotrudniki, u menâ terpeniâ ne hvataet i â obyčno prosto zaglâduû k nim čerez RAdmin... a vot kogda mne dolgo pišut te u kogo â ne âvlâûs′ adminom i â nemogu posmotret′ čto tam delaetsâ... scuko... â nervničaû...))

Shit... in the 4th year of my job as a system administrator I've developed a complex... when somebody has spent a long time writing me a message... I mean, when ICQ is showing that somebody is writing to me, so when colleagues have been spending a long time writing to me, I don't have enough patience, and I usually just take a peek at their screens through RAdmin...[52] but when somebody for whom I am not a system administrator has been spent a long time writing me a message and I cannot see what's happening there... bitch... then I'm nervous...))

In example (6), the speaker confesses that they dislike gaps in ICQ conversations, at least when they are the one waiting. Since their administrator position enables them to satisfy their curiosity at any moment, they have become used to this affordance, and suffer when it is not available. It is this suffering that creates much of the comic effect, since being unable to access the message-in-progress of the interlocutor is an inherent feature of QS-CMC which affects all its users.

Example (5) is more ambiguous. It is tempting to interpret it as an acknowledgment of how important it is not to type anything while your interlocutor is typing to you, i.e. not to "interrupt". This interpretation provides strong support for the claim, made in the analysis of example (4), that such actions can indeed sometimes be viewed by

[52] A remote control software, which allows the administrator to access any machine where it is installed and to see its desktop as users see it (including any text they can be typing).

the speakers as interruptions. The understanding of “the ability to listen” as the ability not to interrupt is quite reasonable.

It is, however, possible, that “the ability to listen” is understood by the speaker as the ability to tolerate gaps, not to avoid overlaps. Indeed, as we see from examples (1) and (6), gaps can be unpleasant for those who are waiting for a message to arrive, and they can manifest that impatience explicitly. Finally, it is also possible that the author of the maxim in example (5) did not intend either of these specific meanings and simply wanted to observe something along the lines of “previously listeners listened to the speakers, now they sometimes have to wait and look at an info message before they can actually start listening”.

3.2. Different speeds of typing

Another affordance is created by the difference between the users in the speed of typing, which can be quite large. The affordance can manifest itself in other types of CMC as well, but most probably not in the same way as in quasi-synchronous communication. Example (7) shows how superiority in the speed of typing can be used for a putdown.

(7) (2012)

1	xxx: вот я и говорю что ты лох
2	xxx: и олень
3	xxx: и ишак
4	xxx: и медленно пишешь
5	xxx: и очень медленно пишешь
6	xxx: и олень
7	yyy: *FUCK*
8	xxx:)))))

1	xxx: vot â i govorû čto ty loh
2	xxx: i olen′
3	xxx: i išak
4	xxx: i medlenno pišeš′
5	xxx: i očen′ medlenno pišeš′
6	xxx: i olen′
7	YYY: *FUCK*
8	xxx:)))))

1	xxx: that's what I'm saying, you're a sucker
2	xxx: and a dummy
3	xxx: and a dumbass
4	xxx: and slow at writing
5	xxx: and very slow at writing
6	xxx: and a dummy
7	YYY: *FUCK*[53]
8	xxx: :::::

In this extract, xxx begins by performing a very straightforward putdown of YYY, i.e. by calling him names (1–3). After receiving no response, xxx attempts to offend YYY by calling him a slow writer (4–5), and then reverts to name-calling (6). The repeat of *dummy* (6) can be viewed as reinforcement of utterances (4–5): "you are so slow I do not even have to invent new insults, just repeating the old ones will do". Most probably, all the time up to YYY's response (7), xxx was seeing the info message which confirmed that YYY was indeed typing a message, not just being totally absent.

The late response can be explained by one (or several) of the following reasons: YYY might indeed be much slower at typing than xxx; he might be crafting his reply; he might be stopping to read xxx's messages (and thinking about responses to them); he might be temporarily distracted etc. In any case, the conversation looks exactly as xxx suggests: that YYY is (relatively) slow. As a consequence, xxx has the floor to

[53] Note that *FUCK* is not a translation, but what YYY actually typed. Expressions of the form *WORD* often mean 'My feelings/thoughts could be conveyed by an emoticon with an html-code WORD (but this emoticon does not exist, or is unavailable to me, or I do not want use it)'.

himself and thus gains a certain dominance in the conversation. This situation can arise in a FtF conversation as well, but if it is achieved by the same means as in example (7), i.e. if Alice is speaking and producing several utterances while Bob is still struggling with the production of his first utterance, the effect would be different. In this case, Alice would be interrupting Bob, not allowing him to speak normally. In example (7), thanks to quasi-synchronicity, xxx is not creating any objective obstacles for YYY, just using the gaps that have emerged due to the difference in typing speed.

The observation that high typing speed can serve as a certain advantage in verbal battles is supported by example (8).

(8) (2011)

1	24: Я никогда, никогда не спорю с ним в ICQ...
2	12: Почему?
3	24: У него скорость за полтысячи знаков в минуту!
4	12: И что???
5	14:[54] А то, что бесполезно спорить с человеком, который пишет быстее, чем ты читаешь.

1	24: Â nikogda, nikogda ne sporû s nim v ICQ...
2	12: Počemu?
3	24: U nego skorost′ za poltysâči znakov v minutu!
4	12: I čto???
5	14: A to, čto bespolezno sporit′ s čelovekom, kotoryj pišet bystee, čem ty čitaeš′.

1	24: I never ever argue with him in ICQ...
2	12: Why?
3	24: He types at more than five hundreds characters a minute!
4	12: So what???
5	14: It makes no sense arguing with a person who can type faster than you can read

[54] I do not know whether this was genuinely said by a user 14 or whether it was actually said by user 24. The numbers most likely were inserted by a poster in lieu of real nicknames, and *14* is a plausible typo.

3.3. Quasi-synchronicity per se

Example (9) is superficially similar to examples (2) and (3), since one of the interlocutors sends a message before the message that could have triggered it arrives. In this case, however, it occurs unintentionally, spoiling the pre-planned teasing sequence.

(9) (2010)

1	Лайк (20:05:21 13/08/2010): кем ты работаешь?
2	Анна (20:44:54 13/08/2010): сейчас - администрирую)))
3	Лайк (20:45:23 13/08/2010): чего администрируешь?
4	Лайк (20:46:18 13/08/2010): Ну ты лошара
5	Лайк (20:46:20 13/08/2010): эээ
6	Лайк (20:46:27 13/08/2010): поспешил
7	Лайк (20:46:35 13/08/2010): так чего ты администрируешь?

1	Lajk (20:05:21): kem ty rabotaeš′?
2	Anna (20:44:54): sejčas - administriruû)))
3	Lajk (20:45:23): čego administrirueš′?
4	Lajk (20:46:18): Nu ty lošara
5	Lajk (20:46:20): èèè
6	Lajk (20:46:27): pospešil
7	Lajk (20:46:35): tak čego ty administrirueš′?

1	Lajk (20:05:21): what do you do?
2	Anna (20:44:54): currently I’m a sysadmin)))
3	Lajk (20:45:23): what kind of network network?
4	Lajk (20:46:18): you loser
5	Lajk (20:46:20): oops
6	Lajk (20:46:27): a bit too quick
7	Lajk (20:46:35): so what kind of networkdo you administrate?

Having asked (3) what the network Anna is administrating is like (i.e. how large it is, which operating system it uses etc.), Lajk must have at once begun typing a predetermined reaction to her response. No matter what she said, he was going to

classify her as a loser. The message-in-progress, however, gets sent too early, thus revealing Lajk's plans to Anna and spoiling the teasing sequence he prepared.

3.4. Encoding issues

There are several encodings which make it possible to represent Cyrillic characters. This is a potential source of problems for speakers of languages using Cyrillic scripts (Paulsen 2011). If encodings used by two interlocutors differ, or if a server transmitting the message does not recognize the encoding correctly, the message can become unreadable. For instance, the situation that occurs in example (10) is possible: yyy is receiving xxx's messages in broken encoding, while yyy's messages reach xxx in a correct form, and thus xxx is not aware that there is a problem. In this case, broken encoding results in an almost total loss of information: every Cyrillic character in xxx's messages to yyy is replaced by a question mark. Thus, yyy cannot attempt to remedy the situation by simply changing their encoding settings or somehow deciphering the messages. They can, however, calculate how many words each message contains and how long these words are.

(10) (2008)

1	xxx: ?????? :)
2	yyy: Прив
3	xxx: ??? ?????
4	yyy: Норма, сам как?
5	xxx: ???? ??????...????? ???? ?? ??????
6	yyy: ХЗ, может пивка попьм??
7	xxx: ??????
8	yyy: Зае**л у тебя с кодировкой хня какая-то я не могу все твои фразы понимать.))

1	xxx: ?????? :)
2	yyy: Priv
3	xxx: ??? ?????
4	yyy: Norma, sam kak?
5	xxx: ???? ??????...????? ???? ?? ??????
6	yyy: HZ, možet pivka pop′m??
7	xxx: ??????
8	yyy: Zae**l u tebâ s kodirovkoj hnâ kakaâ-to â ne mogu vse tvoi frazy ponimat′.))

1	xxx: ?????? :)
2	yyy: Hi
3	xxx: ??? ?????
4	yyy: Fine, you?
5	xxx: ???? ??????...????? ???? ?? ??????
6	yyy: dunno, fancy a beer??
7	xxx: ??????
8	yyy: f*ck off you're having encoding problems and I can't understand every sentence you type. :))

From the number of letters in the first message of xxx (1), yyy manages to guess that he must be saying *Privet :)* 'Hello :)', i.e. a typical conversation opening. Seeing the possibility to engage in a play, yyy does not inform xxx about the encoding problem, but answers with a second-pair part (2). The next adjacency pair is also quite predictable, and yyy assumes that the message in line (3), which looks like *??? ?????*, actually says *Kak dela?* ('How are you?'). Again, yyy responds with a suitable second-pair part.

At this point the opening of the conversation, with its predefined adjacency pairs, ends, and the next message is less predictable. It also turns out to be rather long and complex (5). The response of yyy (6) can be seen as pursuing the following strategy: if xxx is asking something along the lines of "what are your plans for tonight", then my response will suit perfectly; if xxx is asking something else, then my *dunno* can be viewed as the answer to his question, and the offer of a beer will switch the topic of conversation.

The response of xxx (7), however, cannot be deciphered and does not even allow yyy to understand whether their strategy has succeeded (xxx might not have been asking a question at all in (5), which would make *dunno* look strange; or it might have been a question for which *dunno* is not a suitable answer etc.). In (8), yyy gives up.

Note that this playful conversation becomes possible due to unique combinations of factors: messages are being delivered, but are unreadable (although the exact number

of symbols in each word can be calculated), and the sender is not aware of this situation until he is explicitly informed of it.

3.5. Ban

In multi-user chats, there is often an affordance of banning a user who violates certain rules (this can be viewed as an affordance for moderators, but as a constraint for those being banned or those modifying their actions in order not to be banned. In any case, it can be said that the environment provides a certain opportunity for action). A ban can take various forms: in the strictest version, a user is expelled from a chat and is not allowed to rejoin; in the mildest version, as in example (11), a user is temporarily denied the right to send messages, but can still be logged on to chat and read what others are writing (this is important for the interpretation of the example). In Russian-language virtual space, a common reason for being banned is swearing. A bot sometimes monitors conversations and bans users who use profanities included in a predefined list. This is what happens in example (11).

(11) (2010)

1	xxx: Давай вирт?
2	yyy: пошол нахуй урод!!!
3	*Бан: yyy, 10 Минут | Мат.
4	xxx: А теперь, когда ты беззащитна, я буду тебя трахать!
5	xxx: Я медленно снимаю с тебя футболку...

1	xxx: Davaj virt?
2	yyy: pošol nahuj urod!!!
3	*Ban: yyy, 10 Minut | Mat.
4	xxx: A teper′, kogda ty bezzaŝitna, â budu tebâ trahat′!
5	xxx: Â medlenno snimaû s tebâ futbolku...

1	xxx: Wanna have virtual sex?
2	yyy: fuck off, you creep!!!
3	*Ban: yyy, 10 minutes | Swearing.
4	xxx: And now that you are helpless I'm going to shag you!
5	xxx: I'm slowly taking off your T-shirt...

In line (1), xxx invites yyy to engage in virtual sex, a popular activity in CMC since its early days (Deuel 1996). Finding his offer inappropriate, she rejects it (2), insulting xxx and using profanities. For swearing, she is banned, most likely by a bot, although bans are also issued by human moderators. The fact that yyy is banned is made known to xxx (and other participants, if they are present) by means of a service message (3). On receiving that message, xxx claims he will still have virtual sex with yyy (4), and begins doing so (5).

Importantly, what happens in line (5) and probably afterwards can indeed be viewed as virtual sex between xxx and yyy, with yyy unable to protect herself. In the space of a role-play game, where all the actions are performed with words, any verb can be used as an Austinian performative.[55] Virtual sex is that type of game, and its conventions are well-known to chat users, so xxx reasonably views his words as performatives and enjoys his dominance over yyy, who is unable to speak and thus to act. In a normal situation, yyy would be able to respond (i.e. to say "I'm fighting you", if she is willing to support the playful mood to some extent, or to repeat that she does not want to engage in virtual sex, etc.). Being banned, however, she is unable to resist verbally, and the only thing she can do is to log off (if she did so, xxx most likely stopped his actions immediately, since they required yyy's co-presence to be viewed as "real"). The inability of yyy to contradict xxx is somewhat similar to the position of YYY in example (7), although it occurs for different reasons and must be perceived differently.

[55] I owe this observation to Ivan Derzhanski (2003, p.c.)

Why is this situation funny? Again, because it is similar to its prototype, to a situation that could have occurred in real life, but still is obviously very different from it, and this partial similarity (coupled with the fact that no physical harm was done to anybody) creates the comic effect.

3.6. Persistence of transcript

Persistence of transcript (Herring (2007), or conversational persistence (Herring 1999), is what makes QS-CMC less ephemeral than oral speech: messages do not disappear at once, but are stored and visible to participants for at least a certain period of time. Herring (2007) supposes that greater persistence heightens meta-linguistic awareness, and that, in turn, allows users to play with language "in ways that would be difficult in speech". The idea that visible language can increase meta-linguistic awareness and thus stimulate verbal play was also expressed by Werry (1996, 59). Persistence of transcript plays a key role in example (12).

(12) (2012)

1	xxx: вот так разрываются шаблоны.
2	yyy: что?
3	xxx: даже в переписке встречаются глухие.
4	yyy: ты обкурился?
5	xxx: хорошее имя. знаешь с какой фразы нужно начинать деловую беседу?
6	yyy: не знаю
7	xxx: твое любимое женское имя..
8	yyy: русское?
9	xxx: например, какое поле исследовал Фарадей?
10	yyy: а?! мне надоели эти дурацкие вопросы.
11	xxx: ты хоть раз прошел тест на IQ?
12	yyy: Да!
13	xxx: у тебя есть тайные желания? может хочешь чтоб тебя изнасиловали три негра?
14	yyy: хватит строить из себя Фрейда
15	xxx: ладно. прочитай все в обратном порядке =)

1	xxx: vot tak razryvaûtsâ šablony.
2	yyy: čto?
3	xxx: daže v perepiske vstrečaûtsâ gluhie.
4	yyy: ty obkurilsâ?
5	xxx: horošee imâ. znaeš′ s kakoj frazy nužno načinat′ delovuû besedu?
6	yyy: ne znaû
7	xxx: tvoe lûbimoe ženskoe imâ..
8	yyy: russkoe?
9	xxx: naprimer, kakoe pole issledoval Faradej?
10	yyy: a?! mne nadoeli èti durackie voprosy.
11	xxx: ty hot′ raz prošel test na IQ?
12	yyy: Da!
13	xxx: u tebâ est′ tajnye želaniâ? možet hočeš′ čtob tebâ iznasilovali tri negra?
14	yyy: hvatit stroit′ iz sebâ Frejda
15	xxx: ladno. pročitaj vse v obratnom porâdke =)

1	xxx: that’s how expectations are deceived.
2	yyy: what?
3	xxx: even in chat one can meet deaf people.
4	yyy: are you stoned?
5	xxx: nice name. do you know which phrase you should use to begin a business conversation?
6	yyy: I don’t know
7	xxx: your favorite female name..
8	yyy: Russian?
9	xxx: for instance, which field did Faraday study?
10	yyy: what?! I’m fed up with your stupid questions.
11	xxx: have you even so much as once passed an IQ test?
12	yyy: Yes!
13	xxx: do you have secret desires? maybe you want to be raped by three black guys?
14	yyy: stop playing Freud
15	xxx: ok. read everything in reverse order =)

The conversation looks completely disorganized, xxx’s messages not being appropriate second pair parts for yyy’s questions or statements, until xxx’s final remark in line (15). Reversing the conversation (see 12a) and reading it as if it were real, we can now see that xxx and yyy’s messages, starting from line (3) form perfect adjacency pairs.

All yyy’s answers are funny and make him look stupid. He answers “yes” to the question about whether he wants to be raped by three black guys (3-4), he can be said not to know what an IQ test is (5-6), he assumes that Faraday studied the Russian (and not the electromagnetic) field (7-8), he claims “I don’t know” to be his favorite

female name (9-10), which is explicitly ridiculed by xxx in line (11), he supposes that "are you stoned?" is an appropriate opening for a business conversation (11-12), and finally, he responds "what?" to xxx's statement about the existence of deaf people in chat (13-14).

(12a)

1	xxx: ok. read everything in reverse order =)
2	yyy: stop playing Freud
3	xxx: do you have secret desires? maybe you want to be raped by three black guys?
4	yyy: Yes!
5	xxx: have you even so much as once passed an IQ test?
6	yyy: what?! I'm fed up with your stupid questions.
7	xxx: for instance, which field did Faraday study?
8	yyy: Russian?[56]
9	xxx: your favourite female name..
10	yyy: I don't know
11	xxx: nice name. do you know which phrase you should use to begin a business conversation?
12	yyy: are you stoned?
13	xxx: even in chat one can meet deaf people.
14	yyy: what?
15	xxx: that's how expectations are deceived.

Unlike other examples that look as if they occurred spontaneously, in some cases due to a lucky combination of circumstances, this one is obviously a set piece that xxx must have prepared in advance (or learned about from somewhere).

The playful device itself (the incongruity created by an inconsistent text being resolved in the last line by advising the reader to read in a reverse order) might not be new (it is, for instance, used in a poem *Забота* / *Zabota* 'Care' by Genrietta Lâhovickaâ, http://liakhovitskaia.gugunet.de/datein/detjam_4udesjata.html#6[57]).
However, the sophisticated *conversational* teasing of yyy could not have occurred and been understood without persistence of transcript.

[56] Russian field (russkoe pole) is a collocation in Russian, a symbol of Russia's vast spaces, and a name of a well-known folk song.

[57] I am grateful to Olga Podlesskaâ for bringing this poem into my attention.

4. Conclusion

In the preceding section, I have shown how speakers use the affordances of QS-CMC for the amusement of themselves or their interlocutors. The affordances analyzed include quasi-synchronicity per se; availability of info messages and differences in the speed of typing (both inherently related to the quasi-synchronous nature of communication); issues with the encoding of Cyrillic characters[58]; automatic banning of users violating certain rules of conduct; persistence of transcript.

Certain combinations of affordances allow for playful devices that cannot be used in other communication channels, and thus can be considered novel patterns of language play. Some of these patterns achieve a certain measure of popularity. Examples (1), (2), (7) and (10) are not unique in my collection. For every device used in these four examples, one or more other examples are available, either fully isomorphic to the ones reviewed in the article or exhibiting a certain variation. Note that due to the nature of my data (examples collected through not-fully-systematic indirect observation), this fact cannot be viewed as conclusive evidence that these devices are more popular than others. It does, however, provide evidence that the emergence of these devices is not random, and that different speakers invent them, encountering the same affordances. Obviously, this creative use of affordances is limited neither to QS-CMC, nor to conversational structure. My collection includes examples of affordance-dependent play with the relations between *signans* and *signatum*; with obscenity; with double meanings; and with linguistic norms.

In the Introduction, I mentioned the research direction proposed by Sannikov (2002) — use language play to learn more about language, rely on the anomaly to

[58] This is the only affordance that seems to be language-specific, although a conversation similar to example (10) probably can occur in any other language that uses non-Roman script. It was not the aim of this study, however, to compare language play in Russian CMC to language play in other languages.

understand the norm better. Certain conclusions about what is expected in QS-CMC conversations can be drawn from the examples.

As elsewhere, speakers tend to avoid gaps and overlaps, although the latter are not as physically disturbing as in oral speech. Zitzen and Stein (2004, 1000), discussing the conventions used to avoid unexplained gaps in chats, conclude that "there are conversational rules governing behavior with respect to silence which have to be made explicit in order to be taught, as they are not *naturally* acquired". The grounds for this conclusion are unclear to me, and I do not see neither from Zitzen and Stein's nor from my own data in what way the acquisition of conversational rules in chats is less natural than in oral communication. This is, however, an open question, which can be addressed empirically.

Let us return to the list of "lessons" about language drawn from language play. Phatic parts of conversation (openings) are highly formulaic and thus easily predictable (and, in fact, often omitted in QS-CMC, where the conversation can start without a greeting). Quasi-synchronicity can be used for various (not necessarily playful) actions with one's utterance-in-progress that are unavailable in synchronous communication: it can be edited, deleted or prepared in advance in order to save time (cf. example 9).

How aware are lay speakers that their communication is different to that which occurs in the "traditional" channels, and that it contains certain novel patterns? My examples suggest that the level of meta-linguistic awareness is quite high: speakers explicitly refer to their own or their interlocutors' linguistic actions (1, 2, 4, 7, 9, 10, 11), discuss constraints and affordances (5, 6, 8) and use sophisticated techniques of play (12).

Moreover, in several examples (4, 5, 11) the comic effect is created by the *partial similarity* between a certain pattern in QS-CMC and in prototypical FtF communication, and this similarity is more or less explicitly highlighted. In other words, users are aware that what they are doing (or saying, which, in this

communication channel, is equivalent to doing) is amusingly similar to what they can do "in real life", although not fully identical.

To conclude, I would like to return to the general question mentioned at the beginning of the article. As shown here using the example of language play and quasi-synchronous communication, CMC does affect the linguistic behavior of its users. On the one hand, the examples reviewed are not examples of *language change*, at least not in the strict sense of variationist sociolinguistics. On the other hand, they constitute new patterns of linguistic behavior, increase the variation of this behavior and can thus serve as the precursors of potential future changes.

Acknowledgments

Many people provided valuable comments on the various versions and fragments of this text. I am particularly grateful to Jane Stuart-Smith, Edit Bugge, Paul Kerswill, Ingunn Lunde, Bodil Moss, Martin Paulsen, Aleksandr Piperski, Victoria Rosén, Tatjana Russita, Ellen Rutten and an anonymous reviewer for the *Journal of Computer-Mediated Communication*.

References

Anderson, Jeffrey, Fred Beard and Joseph Walther. 2010. "Turn-Taking and the Local Management of Conversation in a Highly Simultaneous Computer-Mediated Communication System" *Language@Internet* 7, article 7. http://www.languageatinternet.org/articles/2010/2804, accessed January 8, 2013.

Anis, Jacques. 2007. "Neography: Unconventional spelling in French SMS text messages." In *The Multilingual Internet. Language, Culture, and Communication Online*, edited by Brenda Danet and Susan Herring, 87–115. Oxford: OUP.

Baron, Naomi. 2010. “Discourse Structures in Instant Messaging: The Case of Utterance Breaks.” *Language@Internet* 7, article 4. http://www.languageatinternet.org/articles/2010/2651, accessed January 8, 2013.

Baym, Nancy. 1995. “The Performance of Humor in Computer-Mediated Communication.” *Journal of Computer-Mediated Communication* 1 (2).

Bechar-Israeli, Haya. 1995. “From 〈Bonehead〉 TO 〈cLoNehEAd: nicknames, play, and identity on Internet Relay Chat”. *Journal of Computer-Mediated Communication* 1(2).

Beißwenger, Michael. 2008. “Situated Chat Analysis as a Window to the User’s Perspective: Aspects of Temporal and Sequential Organization.” *Language@Internet* 5, article 6. http://www.languageatinternet.org/articles/2008/1532, accessed January 23, 2013.

Belz, Julie and Jonathon Reinhardt. 2004). “Aspects of advanced foreign language proficiency: Internet-mediated German language play”. *International Journal of Applied Linguistics* 14(3): 324–362.

Berglund, Therese. 2009. “Disrupted turn adjacency and coherence maintenance in instant messaging conversations”. *Language@Internet*, 6, article 2. http://www.languageatinternet.de/articles/2009/2106, last accessed January 8, 2013.

Burkhaller, Byron, Marc Smith, and JJ Cadiz. 2000. Conversation trees and threaded chats. In *CSCW '00 Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 97–105. New York: ACM.

Cherny, Lynn. 1999. *Conversation and community: Chat in a virtual world*. Stanford: Center for the Study of Language and Information.

Chovanec, Jan. 2011. “Humour in quasi-conversations: constructing fun in online sports journalism.” In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 243–264. Amsterdam, Philadelphia: John Benjamins.

Crystal, David. 2006. *Language and the Internet* (2nd ed.). Cambridge: CUP.

Danet, Brenda. 2001. *Cyberpl@y: Communicating Online*. Oxford: Berg.

Danet, B., Ruedenberg-Wright, L. and Rosenbaum-Tamari, Y. (1997). "«HMMM…WHERE'S THAT SMOKE COMING FROM?». Writing, Play and Performance on Internet Relay Chat". *Journal of Computer-Mediated Communication* 2 (4).

Deuel, Nancy. 1996. "Our passionate response to virtual reality." In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 129–146. Amsterdam, Philadelphia: John Benjamins.

Dynel, Marta. 2009. "Beyond a joke: types of conversational humour." *Language and Linguistics Compass* 3 (5): 1284–1299.

Dynel, Marta. 2011a. "Pragmatics and linguistic research into humour." In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 1–15. Amsterdam, Philadelphia: John Benjamins.

Dynel, Marta. 2011b. "Joker in the pack: towards determining the status of humorous framing in conversations." In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 271–242. Amsterdam, Philadelphia: John Benjamins.

Garcia, Angela, and Jennifer Jacobs. 1999. "The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication." *Research on Language & Social Interaction* 32 (4): 337–367.

González-Lloret, Marta. 2011. "Conversation Analysis of Computer-Mediated Communication." *CALICO Journal* 28 (2): 308–325.

Goodwin, Charles and John Heritage. 1990. "Conversation Analysis." *Annual Review of Anthropology* 19: 283–307.

Gottlieb, Nanette. 2010. “Playing with Language in E-Japan: Old Wine in New Bottles.” *Japanese Studies* 30 (3): 393–407.

Hård af Segerstad, Ylva. 2002. “Use and Adaptation of Written Language to the Conditions of Computer-Mediated Communication.” PhD diss., University of Gothenburg.

Herring, Susan. 1999. “Interactional coherence in CMC.” *Journal of Computer-Mediated Communication* 4 (4).

Herring, Susan. 2007. “A Faceted Classification Scheme for Computer-Mediated Discourse.” *Language@Internet*, 4, article 1. http://www.languageatinternet.org/articles/2007/761, accessed January 8, 2013.

Hutchby, Ian. 2001. *Conversation and Technology: From the Telephone to the Internet*. Cambridge: Polity.

Lee, Carmen K.-M. 2007.” Affordances and Text-Making Practices in Online Instant Messaging.” *Written Communication* 24: 223–249.

Mechkovskaia, Nina. 2006. “Estestvennyk iazyk i metaiazykovaâ refleksiâ v vek Interneta [Natural language and metalinguistic reflection in the era of the Internet].” *Russkij âzyk v nauchnom osveŝenii*, 12 (2): 165–185.

Morkes, John, Hadyn Kernal, and Clifford Nass. 1999. “Effects of Humor in Task-Oriented Human-Computer Interaction and Computer-Mediated Communication: A Direct Test of SRCT Theory” *Human-Computer Interaction* 14 (4): 395–435.

Norrick, Neal. 2010. “Humor in interaction.” *Language and Linguistics Compass* 4 (4): 232–244.

Paulsen, Martin. 2011. “Digital determinism: the Cyrillic Alphabet in the Age of New Technology.” *Russian Language Journal* 61: 119–142.

Plester, Beverly, and Clare Wood. 2009. “Exploring Relationships Between Traditional and New Media Literacies: British Preteen Texters at School.” *Journal of Computer-Mediated Communication* 14(4): 1108–1129.

Popescu, Carmen. 2011. “Understanding ethnic humour in Romanian jokes”. In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 173–190. Amsterdam, Philadelphia: John Benjamins.

Rouzie, Albert. 2001. “Conversation and Carrying-On: Play, Conflict, and Serio-Ludic Discourse in Synchronous Computer Conferencing.” *College Composition and Communication* 53 (2): 251–299.

Sacks, Harvey. 1974. “An analysis of the course of a joke’s telling in conversation.” In *Explorations in the Ethnography of Speaking*, 337–353, edited by Richard Bauman and Joel Sherzer. Cambridge: CUP.

Sannikov, Vladimir. 2002. *Russkij âzyk v zerkale âzykovoj igry* [Russian language in the mirror of language play] (2nd ed.). Moscow: Âzyki slavânskoj kul′tury.

Simpson, James. 2005. “Conversational floors in synchronous text-based CMC discourse.” *Discourse Studies* 7 (3): 337-361.

Smith, Bryan and Greta Gorsuch. 2004. “Synchronous computer mediated communication captured by usability lab technologies: New interpretations.” *System* 32: 553–575.

Su, Hsi-Yao. 2003. “The Multilingual and Multi-Orthographic Taiwan-Based Internet: Creative Uses of Writing Systems on College-Affiliated BBSs.” *Journal of Computer-Mediated Communication* 9 (1).

ten Have, Paul. 1986. “Methodological issues in conversation analysis.” Bulletin de Méthodologie Sociologique, 27: 23–51. Accessed at http://www.paultenhave.nl/mica.htm#N_1_, January 8, 2013.

Thielemann, Nadine. 2011. “Displays of “new” gender arrangements in Russian jokes.” In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 147–172. Amsterdam, Philadelphia: John Benjamins.

Vandergriff, Ilona, and Carolin Fuchs. 2009. “Does CMC Promote Language Play? Exploring Humor in Two Modalities.” *CALICO Journal* 27 (1).

Warner, Chantelle. 2004. “It’s just a game, right? Types of play in foreign language CMC.” *Language Learning and Technology* 8 (2): 69–87.

Werry, Christopher. 1996. “Linguistic and interactional features of Internet Relay Chat. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 47–61. Amsterdam: Benjamins.

Yates, JoAnne, and Wanda Orlikowski. 1993. “Knee-jerk Anti-LOOPism and other E-mail Phenomena: Oral, Written, and Electronic Patterns in Computer-Mediated Communication.” MIT Sloan School Working Paper, http://ccs.mit.edu/papers/CCSWP150.html

Zitzen, Michaela and Dieter Stein. 2004. “Chat and conversation: a case of transmedial stability?” *Linguistics* 42 (5): 983–1021.

VI

6. Rise and fall of the Olbanian language: diachronic description of an online slang

1. Linguistic background

1.1. Synchrony

One of the numerous possible ways to spell the name of the slang in question is *йазы́к падо́нкафф*[59] / *jazý̀k padónkaff*, while the only standard spelling would be *язы́к подо́нков / âzý̀k podónkov* ('language of scumbags').[60] Let us briefly review the deviations from the standard norm in the former variant to obtain some insight into the essence of this phenomenon.

The letter *я / â* actually denotes a combination of two phonemes, /ja/, so the erratic *йа / ja* is pronounced in the same way. Three out of four deviations in the second word are of a similar type. The pronunciation of the word *подо́нков / podónkov* can be transcribed as [padónkaf].[61] Phonemes /o/ and /a/ are distinct only in stressed position, while in an unstressed syllable both are realized as a reduced vowel, which in the given phonetic contexts is close to [a]. Phoneme /v/, as with all paired voiced consonants, is devoiced at the end of the word and pronounced as [f].

Thus, these deviations follow the phonetic principle "write as you hear". The standard orthography, however, often ignores that principle in favor of a principle which can be labeled as *phonemic* - "write an underlying phoneme" - or *morphemic* - "write in a way that retains the most consistent spelling of morphs across different

[59]All Russian examples are given in the ISO 9 transliteration system (ISO 1995), the only system with biunique correspondence between Cyrillic and Latin characters. Differences between erratic and standard forms are underlined where possible. Stress is always marked (this is not usually done in Russian, apart from in children's books and textbooks).

[60] In this article, I use the names *Olbanian* and *язык подонков* fully interchangeably, although some researchers (e.g. Hristova 2011) endow these names with different meanings.

[61] A simplified transcription is used, IPA would be [pɐdónkəf].

word forms" (on the principles of Russian orthography see Knâzev and Požarickaâ 2012, 376–386).

The remaining erratic spelling, *фф / ff* instead of *ф / f*, is somewhat different. Just like the others, it does not change the pronunciation of the word: double consonants in this position would be indistinguishable from a single one in normal speech. It cannot, however, be said to follow the "write what you hear" principle, rather the opposite one: "write **not** what you hear, but what would sound the same".

This introduction provides some insight into the spirit of the Olbanian language: the violation of linguistic norms, predominantly the norms of orthography, predominantly by following the principle "write as you hear". Some deviations also concern phonology and morphology, but they are marginal compared to those that relate to orthography (see Sebba 2003 for an overview of the cases when violation of orthographic norms manifests a rebellion spirit). Another distinctive feature of Olbanian is the active use of clichés. Some of these clichés contain words that have not existed previously or have acquired new meaning (see Gusejnov 2005 for a brief wordlist with explanations), but for the most part they also depend heavily on anti-orthography to make them visible and distinguishable.

This triggers the question of whether Olbanian is indeed a *slang*. Eble (1996, 11) defines slang as "an ever changing set of colloquial words and phrases that speakers use to establish or reinforce social identity or cohesiveness within a group or with a trend or fashion in society at large". Lighter (1994, xi) offers the following working definition (which, in his own view, is not fully adequate, since it fails to take the social dimension into account): "an informal, nonstandard, nontechnical vocabulary composed chiefly of novel-sounding symbols for standard words and phrases". The latter is rather close to Halliday's (1976) notion of *anti-language*. Note that Halliday (1976, 571) also considers relexicalization ("new words for old") to be a basic means of creating an anti-language, although he acknowledges that foregrounding of certain nonstandard elements also occurs at other levels, such as the phonological, lexicogrammatical and semantic.

In our case, the level which conveys most of the *anti* spirit is orthography. Relexicalizing a word usually means rewriting it using Olbanian orthographical (anti-)conventions. The increased importance of orthography is unsurprising, given the general trend in modern language culture towards both greater autonomy and greater salience for written language (see Chapter 4). Another widely-known idiom that manifests its slang-like nature primarily through the distortion of writing system is leet speak (see, e.g., Blashki and Nichol 2005). Just like Olbanian, it has some distinctive features on other levels, including lexical clichés, and it has also emerged online. In some other respects, however, it is noticeably different. First, the proclaimed intention is different: Olbanian originally positioned itself as an anti-literacy movement (Zvereva 2009), while leet was an elite code. Second, leet replaces Latin letters with other ASCII characters; that is, it affects primarily the graphical level. Olbanian violates the non-trivial sound–letter correspondences; that is, it affects orthography. Third, probably due to these two factors, leet speak turned out to be a less provocative venture. Thus, although it is internationally known, its popularity and salience among English speakers seem to be less than the popularity and salience of Olbanian among Russian speakers.

The role of Olbanian within the Russian-speaking community is quite close to that of slang (as defined above), and so *orthographical slang* (or anti-language) seems to be an appropriate label for it.

1.2. Diachrony

Olbanian has received a fair amount of linguistic attention (Gusejnov 2005, Krongauz 2008, Šapovalova 2008, Zvereva 2009, not to mention numerous articles in the mass media and a self-description (Sokolovskij 2008)). The late Daniela Hristova (2011) offered the most detailed linguistic description to date. When it comes to diachronic descriptions, however, what is available are mostly speculations.

Studies of slang diachrony have never been frequent, and this is understandable, given that slangs usually do not "fossilize", or leave a written record, at least not one that is easily accessible. Within CMC, the situation with data availability is much better, but still, as Androutsopoulos (2011, 150) notes, "...there is a striking lack of systematic micro-diachronic studies within CMC. While the implicit assumption seems to be that digital language innovations are here to stay, 'rise and fall' patterns are just as possible". Some diachronic studies of language in CMC, however, can be mentioned, cf. Herring 1998, Rowe 2011.

Returning to the phrase *йазы́к падо́нкафф / jazýk padónkaff,* we can see that here the anti-normative spirit manifests itself five times in two short words. Producing and perceiving texts with so many deviations can be quite difficult, and Krongauz (2008) assumes that, while the original core principle was to violate the norm as many times as possible, this unreachable standard eroded with time, and it became enough to make at least some errors. This assumption seems plausible and is supported by users' metalinguistic comments, like the following one:

(1)
ORIGINAL EXAMPLE:
бля как чита́ть нудо́бна про́ста пизде́цъ, ме́ньше увлека́йся антиарфагра́фией, падо́нак бля (udaff.com, comment to a text, 2003).

blâ kak čitát' nudóbna prósta pizdéc", mén'še uvlekájsâ antiarfagráfiej, padónak blâ

STANDARD ORTHOGRAPHY[62]:
бля, как читать неудобно, просто пиздец, меньше увлекайся антиорфографией, подонок, бля

blâ, kak čitát' neudóbno, prósto pizdéc, mén'še uvlekájsâ antiorfográfiej, podónok, blâ

TRANSLATION:

[62] Deviations in punctuation marks are also corrected, but are not discussed here: they are not a distinctive feature of Olbanian.

> ‘shit, this is so difficult to read, that’s just fucking shit, use less anti-orthography, you fucking scumbag’

The author of this comment criticizes a fellow “scumbag” for using too much anti-orthography in a text. Note that the comment itself abounds with orthographical deviations (9 per 11 words).

On the other hand, there is no direct psycholinguistic evidence to support the claim that overuse of Olbanian hinders production or perception[63]. Regardless of whether this claim is actually true or not, there is a common view that Olbanian has already passed its peak and in recent years has been dying out, probably as a result of a simple fading of the fashion. This, however, is an intuitive impression that has never been empirically verified, and the exact dates of alleged peaks and troughs are unknown, although it is usually assumed that public interest towards the slang peaked in 2004–2006 (Zvereva 2009, 54).

In this chapter, I aim to verify the existing intuitive assumptions about the diachrony of the Olbanian language, both within the community of its active users (experiment 1) and outside it (experiment 2), thus extending our understanding of how an online slang can develop. I will focus on the behavior of one parameter that provides crucial information about the status of the slang and is easily measurable — the frequency of actual usage of the slang’s distinctive linguistic features. In the case of Olbanian, this means the frequency of norm violations, predominantly of erratic spellings.

[63] Moreover, experiments carried out with leet words (Perea, Duñabeitia, and Carreiras 2008) suggest that the cognitive effort necessary to read them is small. On the other hand, recognizing separate words and reading or writing texts are very different tasks, and the latter has not been studied.

2. Experiment 1.

All the data were taken from the website udaff.com (the center of the *padonki* culture and one of the cradles of the Olbanian language), from the section *kreativy* ('creative stories') where users upload their own short stories. This is one of the oldest and most important sections on the website, and its name is a symbol of *padonki* culture. It was chosen as the largest and most diachronically representative collection of texts a) with a large number of erratic spellings; b) written by people who identify themselves as *padonki,* i.e. "native speakers" of Olbanian. It should be kept in mind, however, that these data do not necessarily reflect patterns of use of the Olbanian language elsewhere.

Texts were selected from 975 webpages covering the time period from January 2001 to December 2011. One text was selected randomly from each page (each page contained 50 texts), and a random fragment of 100 words was extracted for analysis. If a text was for some reason not suitable for analysis (e.g. it was shorter than 100 words), another random text was selected.

This resulted in 975 100-word fragments produced by 729 authors (156 authors produced more than one text, the largest number of texts per author was nine, the mean was 1.34). No adjustment was made for the fact that some authors had more than one fragment included in the sample: while this gives their idiolect additional chances to contribute to the observed variation, that must mirror the actual situation. For every word, it was noted how many deviations from the norm it contained. All kinds of deviations were counted, and not all of them are strictly Olbanian. However, the analysis of distribution of deviations across different types (Appendix 1) shows that the number of indisputably non-Olbanian deviations is relatively small and constant and does not distort the general picture.

The coding was carried out by two research assistants (linguists and native speakers of Russian) and myself.

Results are shown in Fig. 1. It can be seen that there is a sharp decrease in frequency from 2001 to 2002 and then a gradual decrease from 2002 to 2011. In 2001, the general trend is to sprinkle one's text with deviations. On average, there are not so many (the median lies rather low down), but some texts contain a fair number (the upper edge of the box is high, and the upper whisker is sky-high). There are, however, no outliers: even the texts saturated with deviations can still be said to fall within the general trend. Starting from 2002, the median lies much lower down, the box size and the whisker range are smaller and continue to decrease over a period of years. Some authors still produce texts rich in deviations, but these texts are rare now, they are outliers, and even their number falls noticeably towards the end of the period.

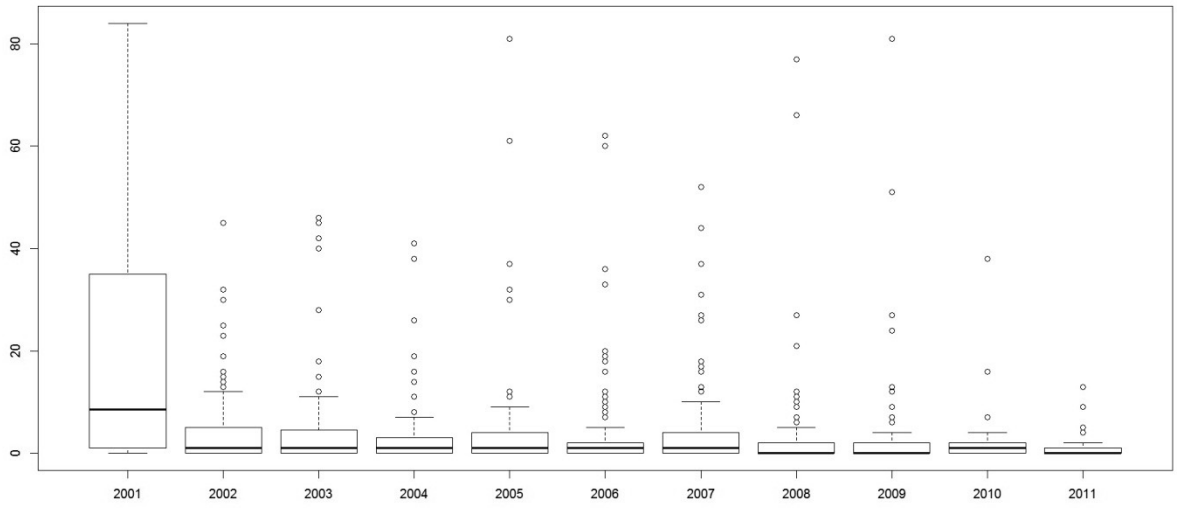

Figure 1. Frequency of use of the Olbanian language on the udaff.com website (*kreativy section)*: number of deviations from the norm per 100 words.[64]

[64] This a boxplot, or a box-and-whiskers plot. Since some readers may be unfamiliar with it, I am providing a brief explanation of what it is and how it should be interpreted. For each year, several dozens of text fragments (from 34 in 2001 to 142 in 2007) were analyzed, which provides us with 11 *distributions* of data points. Each point is a number of deviations in a text fragment. Various strategies can be used to represent a numerical distribution visually. One way is to plot *all* the points. A graph of this kind gives the reader the full information, but would be inconvenient in our case, since too many points make a graph unreadable. Another way is to provide just *one* number to characterize the whole distribution, often the arithmetic mean. A graph of

3. Experiment 2

The data from *kreativy*, used in experiment 1, tell us nothing about the use of the Olbanian language outside the udaff.com community,[65] and it is interesting to know what happens on a broader scale. For that purpose, a "Pulse of the blogosphere"[66] service offered by Yandex, a major Russian search engine, was used.[67] "Pulse" indexes a significant part of the Russian blogosphere and returns a share of all the blog posts that were written within a given month (starting from June 2001) and contain a given search query (that is, a word or a phrase).

To select a number of appropriate queries, the following was done. During experiment 1, a list of all the word forms containing deviations from the norm was compiled. The list contained 3,109 word forms. To constitute a good query, a word had to meet the following criteria: a) an average monthly frequency of occurrence within the blogosphere of at least 100 blog posts (otherwise "Pulse" did not provide reliable results); b) to be the result of using specifically Olbanian slang and not committing a random error or allowing deliberate norm violations for other reasons. To narrow the list down, only forms which had been classified as containing exactly

---

this kind is very simple, but provides no information about how dispersed or how skewed the distribution is. Boxplot uses *five-number* summary to represent a distribution. The black horizontal line is the *median* (half of the observations lie above this point and half lie below, this is the middle of the dataset). The box in the middle indicates the "core" of the distribution (technically speaking, the horizontal edges of the box are hinges, which are almost equivalent to upper and lower quartiles). The length of the box makes it possible to estimate how *dispersed* the distribution is (the longer the box, the higher the dispersion), the position of the median within it indicates whether the distribution is *skewed* (in Fig. 1, for instance, the median lies close to the lower edge, which means that there are more texts with a small number of deviations than with a high number). Whiskers indicate the range within which all the observations are expected to fall (technically, they show the largest or smallest value within a distance that is 1.5 times the box size from the nearest hinge). Dots that lie beyond the whiskers are data points considered *outliers*, observations that strongly deviate from other members of the sample. Informally speaking, these are the points which can be considered "unusual".

[65] Moreover, even in the other sections of the website, the diachronic patterns may be different.

[66] http://blogs.yandex.ru/pulse/

[67] Standard Internet search engines are notoriously imprecise with year-specific search queries, but since "Pulse" was created especially for diachronic studies, the hope is that it can be more reliable.

three deviations from the norm were selected (those which contained more were likely to violate condition a, those which contained less were likely to violate condition b). The remaining forms were manually checked to satisfy conditions a and b. This procedure resulted in ten words: *исчо́ / iščó* ‘more, yet’ (from *ещё / eŝé*), *ниибе́т / niibét* ‘I am not concerned’ (from *не ебёт / ne ebët*[68] literally ‘this does not fuck [me]’), *канéшна / kanéšna* ‘of course’ (from *конéчно / konéčno*), *што́ле / štóle* ‘or what?’ (from *что ли / čtó li*), *йопт*, profane interjection (derived in a somewhat unclear way from the verb *ебать* ‘to fuck’), *нимагу́ / nimagú* ‘I cannot’ (from *не могу / ne mogú*), *апсте́ну / apsténu* ‘against the wall’ (from *об сте́ну*), *ниибáца / niibáca*, *ниибáцца / niibácca* ‘very’ (from *не ебáться / ne ebát'sâ*, literally ‘not to fuck’), *беспезды́ / bespezdý* ‘certainly, honestly’ (from *без пизды́ / bez pizdý*, literally ‘without a cunt’). Some of these words are parts of frequent clichés (e.g. *убúться апсте́ну / ubít'sâ apsténu* ‘kill oneself against the wall’), but this fact was not taken into account. I assume that the frequency of use of these ten words can serve as a proxy for estimating the general popularity of the Olbanian slang.

Results are shown in Fig. 2. The pattern is noticeably different from that observed in Fig. 1: here, the frequency peaks in 2006 with a gradual increase in preceding years and gradual decrease (nearly to extinction) in subsequent years.

To test whether the observed differences are indeed due to the different data source (blogosphere vs. udaff.com) and not the measuring method (number of deviations in text fragments vs. frequency of a few erratically spelled words), the same method was applied to all the *kreativy* available on udaff.com (i.e. the population from which the data for experiment 1 were sampled). Results are shown in Fig. 3. The pattern is very similar to that observed in Fig. 1, although the decrease after 2001 is more gradual. Thus, the results of experiment 1 are confirmed.[69]

---

[68] The spelling of *e* instead of *ë* is allowed by the norm.

[69] Note that Figures 1, 2, and 3 all measure different things, so one cannot make any meaningful direct comparison of absolute numbers provided in them. It only makes sense to compare diachronic trends.

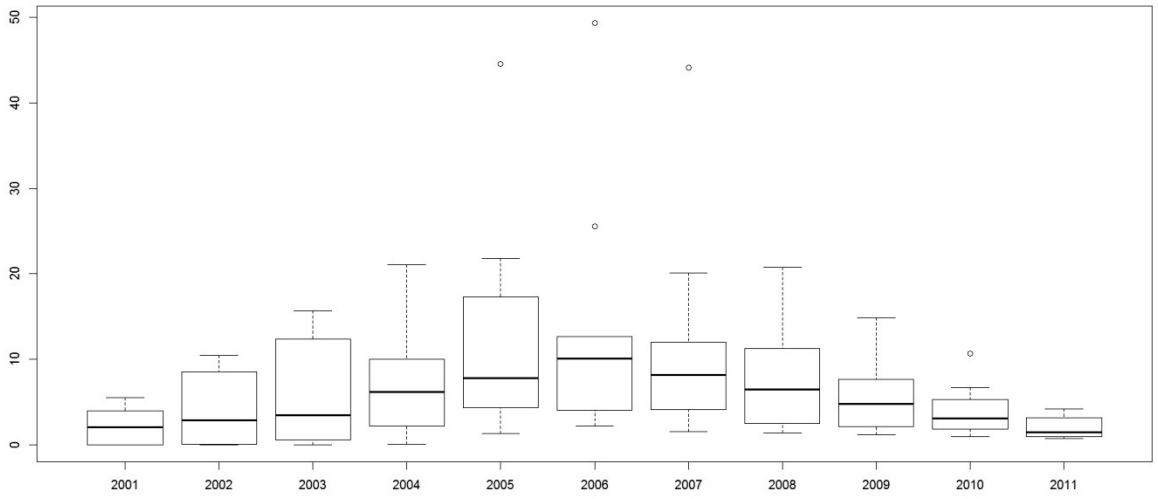

Figure 2. Frequency of use of the Olbanian language in the Russian blogosphere: ratio of the number of blog posts where a query word occurs to the total number of blog posts in the given year; per mille. For each year, the frequency was calculated as the arithmetic mean of frequencies for all 12 months (apart from 2001, for which the data are available starting from June). The query words were the ten words from the selected set.

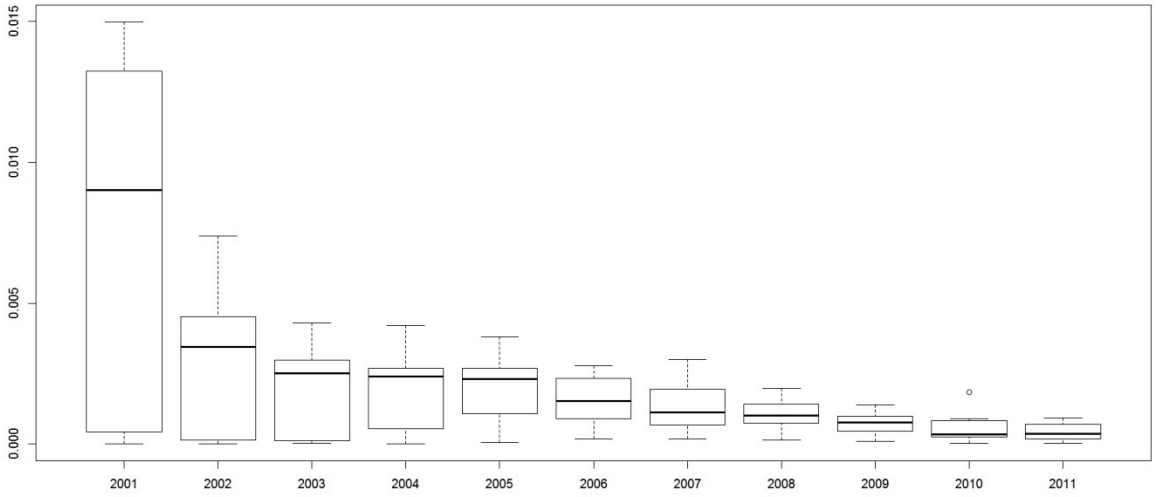

Figure 3. Frequency of use of the Olbanian language on the udaff.com website (kreativy section): weighted frequency of the ten selected words, calculated as the mean number of their occurrences divided by the size of the text file containing all the kreativy for the given year.

4. Discussion

The common view that the Olbanian language has fallen out of fashion is confirmed by both experiments: in 2011, the frequency of deliberate norm violations falls to almost zero. It is, however, unlikely that the decrease in frequency after 2001, observed in Figures 1 and 3, is a result of the same extinction process. In 2002, Olbanian is still a fashionable new practice. The more plausible explanation for this decline is offered by Krongauz's (2008) insight that, after toying extensively with new anti-orthographical devices, speakers start using them more economically and in a more symbolic way. It is somewhat surprising that this happens so early, just a year after the udaff.com website was created and the slang began gaining prominence.

Zvereva's (2009, 54) observation that the public interest towards Olbanian peaked in 2004–2006 fits the results of experiment 2, although Fig. 2. suggests that the heightened interest lasted for a longer period, 2004–2008, with a peak in 2006. In any case, it is noteworthy that speakers *inside* and *outside* the Olbanian community use anti-normative devices most actively in different periods. Interestingly, the absolute number of texts in the *kreativy* section of the udaff.com website also undergoes a rather gradual increase and decrease, with a peak in 2007 (see Fig. 4). It is, difficult, however, to tell whether this is the consequence of the increased public interest or the cause. In any case, at udaff.com, there is no upsurge in the frequency of deviations from the norm.

The early-year decrease is less steep in Fig. 3 than in Fig. 1. This can probably be attributed to the method of measurement: as I mentioned earlier, some of the ten words selected are part of oft-used clichés, and thus their frequency can fall in a more gradual way and exhibit less variation. Interestingly, at the end, it also approaches zero. Thus, the data do not support the common view that Olbanian, while nearly extinct, has been survived by a number of clichés.

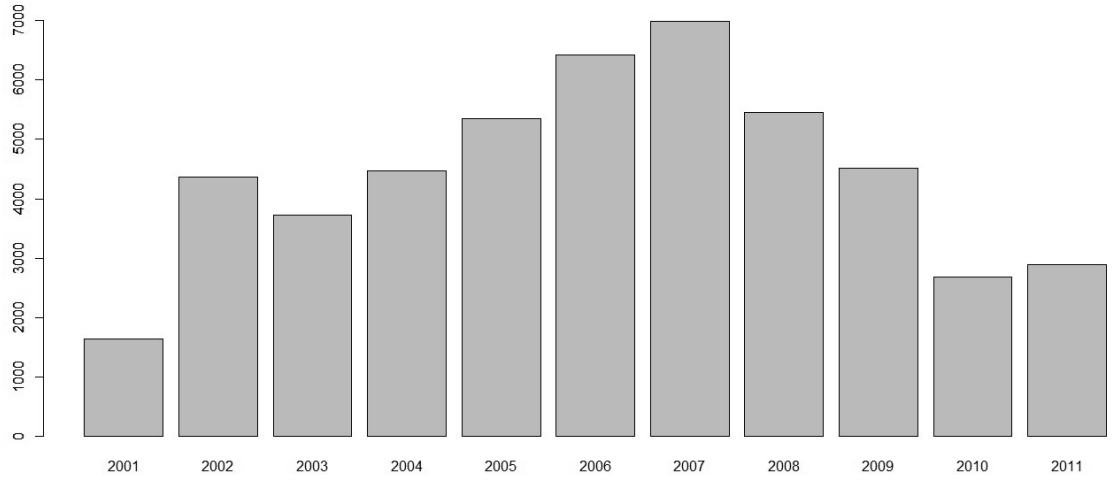

Figure 4. Number of texts per year in the *kreativy* section of the udaff.com website.

5. Conclusion

As I mentioned above, there is very little systematic research on slang diachrony. I am aware of no data that would make it possible to carry out a comparison of Olbanian with another online slang, or, more interestingly, with an offline one. The latter would have made it possible to address the general question of "are there any differences between development patterns of online and offline slangs?" (which can be considered part of the general question of this dissertation "how does CMC influence the linguistic behavior of speakers?").

In the absence of such data (but with the hope that someday they will be accumulated, at least for relatively accessible online slangs), only some preliminary observations can be made. First, as has already been mentioned, we witness the increased role of writing and orthography. It was the denial of orthography and its norms that gave birth to the Olbanian language and that has been the main source of relexicalization. Interestingly, orthography is also a convenient resource for *overlexicalization*, which, according to Halliday (1976, 571), is a typical feature of anti-languages: for one word from the standard language, we often find dozens of

slang synonyms. In Olbanian, one might choose between *язы́к падо́нков / âzýk padónkov*, *йазы́г падо́нкафф / jazýg padónkaff, езы́к подо́нкаф / ezýk podónkaf* and dozens of other possible combinations of deviations from the norm, cf. also *ниибáца / niibáca* and *ниибáцца / niibácca* from experiment 2. Speakers might use this to display their linguistic creativity or to choose how far from the norm they want to deviate. It might be a worthwhile research agenda to study the diachronic behavior of this variation: does it increase with time? Or is the reverse true, and most speakers converge on a certain spelling, thus establishing a proto-norm?

Second, the online processes happen with remarkable speed. Again, in the absence of reliable data it is impossible to claim that they are slower in offline slangs, but it is still impressive that an idiom can become immensely popular and then go out of fashion in the space of some 11–12 years.[70]

This “rise-and-fall pattern”, to use Androutsopoulos’s (2011, 150) apt description, is nicely illustrated by the following extract, where an anonymous user attempts to greet others in the Olbanian way (consider the density of deviations, and the question about “chicks” is also very typical). This, however, happens in 2012, and the speaker makes himself look like a living fossil:

(2)
ORIGINAL EXAMPLE:
xxx: Преве́д, паддо́нке! Каг ди́ла? Ба́бы даю́д?
yyy: Пффф... Тебя́ то́лько разморо́зили что́-ли? (bash.org.ru, 2012)

xxx: Prevéd, paddónke! Kag dilá? Báby daûd?
yyy: Pfff... Tebấ tólʹko razmorózili čtó-li?

STANDARD ORTHOGRAPHY:
xxx: Привет, подонки! Как дела? Бабы дают?
yyy: Пффф... Тебя только разморозили, что ли? (bash.org.ru, 2012)

[70] The history of Olbanian actually begins earlier than 2001, in the 1990s, see Zvereva 2009.

xxx: Privét, podónki! Kak delá? Báby daût?
yyy: Pfff... Tebâ tól′ko razmorózili čtó-li?

TRANSLATION:
xxx: Hello, scumbags! How are you doing? Are the chicks giving it up?
yyy: Phew... Have you just been unfrozen, or what?

Olbanian, however, has left its imprint on the Russian language. Its traces can be found, for instance, in speakers' attitude towards the norm (Zvereva 2009) and in formal linguistic structure (see Chapter 3).

Acknowledgments

I am grateful to Aleksandr Piperski and Liudmila Zaydelman for doing most of the coding-and-counting work.

References

Androutsopoulos, Jannis. 2011. "Language change and digital media: A review of conceptions and evidence". In *Standard languages and language standards in a changing Europe*, edited by Kristiansen Tore and Nikolas Coupland, 145–161. Oslo: Novus.

Blashki, Katherine, and Sophie Nichol. 2005. Game Geek's Goss: Linguistic Creativity In Young Males Within An Online University Forum (94/\/\3 933k'5 9055oneone). *Australian Journal of Emerging Technologies and Society* 3 (2): 77–86.

Eble, Connie. 1996. *Slang and sociability: in-group language among college students.* Chapel Hill: UNC Press.

Gusejnov, Gasan. 2005. “Berloga vebloga: Vvedenie v èrratičeskuû semantiku.” [The lair of the weblog: Introduction into the erratic semantics] http://www.speakrus.ru/gg/microprosa_erratica-1.htm

Halliday, M.A.K. 1976. “Anti-Languages.” *American Anthropologist* 78(3): 570–584.

Herring, Susan. 1998. “Le style du courrier électronique: variabilité et changement.” *Terminogramme* 84-85: 9-16. English version accessed 23 January 2013 at http://ella.slis.indiana.edu/~herring/term.html.

Hristova, Daniela. 2011. “Velikij i mogučij olbanskij âzyk: The Russian Internet and the Russian Language.” *Russian Language Journal* 143–162.

Knâzev, Sergej, and Sof̓â Požarickaâ. 2012. *Sovremennyj russkij literaturnyj âzyk. Fonetika, orfoepiâ, grafika i orfografiâ.* [Contemporary standard Russian langugage. Phonetics and phonology, orthoepy, graphics and orthography] Moscow: Gaudeamus.

Krongauz, Maksim. 2008. “Utomlennye gramotoj.” [Wearied by literacy] *Novyj mir* 5: 153–159. Accessed at http://magazines.russ.ru/novyi_mi/2008/5/kr11.html

Lighter, Jonathan. 1994. Preface to *Historical Dictionary of American Slang*, xi–xxxix. Oxford: OUP.

Perea, Manuel, Jon Andoni Duñabeitia, and Manuel Carreiras. 2008. “OBSERVATIONS: R34D1NG W0RD5 W1TH NUMB3R5.” *Journal of Experimental Psychology Human Perception and Performance* 34(1): 237–241.

Rowe, Charley. 2011. “*Whatchanade?* Rapid language change in a private email sibling code.” *Language@Internet* 8, article 6. http://www.languageatinternet.org/articles/2011/Rowe, accessed January 23, 2013.

Sebba, Mark. 2003. “Spelling rebellion.” In *Discourse constructions of youth identities*, edited by Jannis Androutsopoulos and Alexandra Georgakopoulou, 151–172. Amsterdam, Philadelphia: Benjamins.

Šapovalova, Nadežda. 2007. “Karnaval′noe obŝenie v Internete (na materiale sajta www.udaff.com).” [Carnivalesque communication in the Internet (using the data of the website www.udaff.com] *Vestnik Čelâbinskogo gosudarstvennogo universiteta* 16(20): 164–167.

Sokolovskij, Dmitrij (2008). *Bibliâ padonkov, ili Učebneg albanskogo âzyka*. Moscow: Folio SP.

Zvereva, Vera. 2009. “Âzyk padonkaf”: Diskussii pol′zovatelej Runeta. In *From Poets to Padonki: Linguistic Authority and Norm Negotiation in Modern Russian Culture*, edited by Ingunn Lunde and Martin Paulsen, 49–79. Bergen: University of Bergen.

6A. Appendix 1. The best way to violate the norm: different types of norm deviations in the Olbanian language

A1. Introduction

In experiment 1, the deviations from the norm were not only counted, but also categorized into several types. This makes it possible to find out what kinds of deviations are most popular within the slang, as well as to verify a specific hypothesis about Olbanian's diachronic development, put forward by Hristova (2011) (see section A3 for the formulation of the hypothesis).

A2. Methods

A2.1. Major types

The two major types are "orthography" and "phonology". By "major" I mean not most frequent, but rather those which seem to be most salient in the minds of users and researchers of Olbanian.

The "orthography" type comprises those erratic spellings which violate orthographical conventions, but do not affect the phonology level, i.e. the wrongly spelt word would still be read correctly. Two basic classes can further be singled out within this type, which might be labeled "transcription" and "anti-transcription".

The "transcription" class implies that the erratic spelling follows the phonetic principle, or "write as you hear", while the orthography requires something else. Consider the erratic spelling *padonkaff* instead of the standard form *podonkov* 'head-GEN.PL'.[71] As explained in section 1.1 of the main text, both deviations from standard spelling (*a* instead of *o* and *f* instead of *v*) mirror the actual pronunciation. The erratic

[71] Glosses are in accordance with the Leipzig Glossing Rules (http://www.eva.mpg.de/lingua/resources/glossing-rules.php).

spelling ignores the phonemic principle of the orthography, but instead approaches phonetics.

Another example is *katorovo* instead of *kotorogo* 'which-GEN.SG.M'. Here again, we find *a* instead of *o*, but note that this happens only once, despite the fact that the third and fourth syllable are also unstressed, and the writer might have used these two positions for erratic spelling as well. In addition, we also see *v* instead of *g*: the requirement of standard orthography to spell [v] as *g* in the adjectival genitive ending *-ogo* (for historical reasons) is not observed.

The "anti-transcription" class is the opposite phenomenon, with the erratic spelling following the principle 'write NOT what you hear'. Consider the erratic spelling *kde* instead of *gde* 'where'. The spelling of *k* instead of *g* does not affect pronunciation: due to assimilation a voiceless consonant before a voiced one becomes voiced as well, so *kde* is still read [gd'e]. Another example is *preved* instead of *privet* 'hello'. As well as /a/ and /o/, /i/ and /e/ are distinct only in a stressed syllable. In an unstressed one, they are reduced to a sound quite close to [i]. In *privet*, the first syllable is not stressed, so writing *e* instead of *i* does not affect pronunciation, but actually goes against the principle "write what you hear": you do not hear [e] there. In the same way, writing *d* instead of *t* in *preved* is the opposite of writing *f* instead of *v* in *padonkaff*: in both cases, the distinction voiced–voiceless is neutralized in final position, but the latter imitates pronunciation, while the former contradicts it. Double *f* instead of a single one in *padonkaff* can also be said to follow the same principle (in normal speech, double [f] at the end of a word sounds like a single [f]).

The description of these two classes can be summarized as follows. Russian orthography quite often fails to correspond to actual pronunciation. Both classes make use of these inconsistencies, but in different ways: the first class by "transcribing" speech, the second one by "hypercorrecting" it.

The "phonology" type is different from the "orthography" type in that the phonology of the word is also changed. Consider *medved* instead of *medved'* 'bear' (the final

consonant is hard instead of being soft), *scuko* instead of *suka* ‘bitch’ (two consonants instead of a single [s] in the beginning; *o* instead of *a* falls into the “anti-transcription” class), *defačku* instead of *devočku* ‘girl-ACC.SG’ (voiceless consonant instead of a voiced one, and not in neutralization position). These deviations will persist if these words are read aloud. There is anecdotal evidence that they do occur in oral speech (Hristova 2011, also personal experience), but there are no reliable corpus data.

As the three examples above show, in deviations of this type a sound changes into a similar sound, and often just one phonological parameter (hardness–softness, voicedness–voicelessness) is being altered.

A2.2. Minor types

The minor types include “portmanteau”, “vernacular”, “common morphology”, “Olbanian morphology”, “typo” and “one word or two words”.

Portmanteaus include erratic forms which are motivated by language play, viz. merging two words into one, e.g. *lizbiânki* — a portmanteau of *lesbiânki* ‘lesbians’ and *lizat'* ‘to lick’ (although note that this particular example can also be viewed as two deviations of “transcription” class).

“Vernacular” refers to the use of orthography to render vernacular or allegro-speech pronunciation, mostly contractions and sound omissions: *pât'sât* instead of *pât'desât* (‘fifty’). Some orthographical pronunciation observations of this kind that were often attested before the emergence of Olbanian and became well-established literary devices (like *čë* instead of *čto* ‘what’) were ignored and not considered deviations.

“Common morphology” includes typical morphological deviations like *evonnyj* instead of *ego* ‘his’, while “Olbanian morphology” includes morphological deviations that emerged due to the influence of Olbanian orthography. This type requires some additional clarifications.

Consider the following examples:

(1)
Olbanian form: *kot-eg-ø*
Standard form: *kot-ik-ø*
tomcat-DIMIN-NOM.SG
'tomcat'

(2)
Olbanian form: *kot-eg-a*
Standard form: *kot-ik-a*
'tomcat-DIMIN-GEN.SG'
'of the tomcat'

Both deviations in the erratic form *koteg* from example (1) fall into the "anti-transcription" class of the "orthography" type. Nominative singular is the only form of this word with zero ending, in all the other forms the suffix [ik] is followed by a vowel, and the opposition [k]-[g] is never neutralized before vowels. In example (2), the speaker ignores this fact in order to follow the phonemic principle of Russian orthography, that is, to keep the *written* paradigm of the word regular. As a result of that, other word forms (in example 2, genitive singular) are also spelt with *g*, and this affects phonology. Krongauz (2008) and Hristova (2011) describe similar cases as the emergence of new suffixes (in this case suffix *-eg* replaces suffix *-ik*), but the same process can actually affect the last consonant of a root: cf. *druk̲om* ('friend-INSTR.SG') instead of a standard *drugom*, derived from erratic nominative *druk̲* ("transcription" class) instead of *drug*.

Hristova (2011) describes several other deviations of this kind. Unlike her, I do not consider the famous erratic spelling of the endings of reflexive verbs — *-c(c)o* or *-c(c)a* instead of *-t′sâ* (infinitive) or *-t′sâ* (3sg present) — to be an example of morphological change. This deviation is not morphologically conditioned by the combination of *-t* or *-t′* and reflexive suffix *-sâ*. Quite the contrary, it is purely phonetic: *-t(')sâ* is pronounced as [ca]. The erratic spelling *c(c)* instead of *mьc* can also occur in a completely different morphological context, cf. *nuц(ц)om/pic(c)ot* instead of *пятьсóт/pât'sót* 'five hundred'. On the other hand, it should be noted that

such combinations are rare, and the erratic pattern *t(')s—>c(c)* is indeed usually associated with the reflexive verbal endings.

The "Olbanian morphology" type is not independent: it occurs only in combination with other types ("phonology" or "orthography"). Nevertheless, it is worthy of attention, since it represents probably the deepest structural influence Olbanian orthography has had on the Russian language. One case of such influence (gender change) is analyzed in detail in (Berdičevskij 2012).

The "one word or two words" type includes erratic spellings achieved by ignoring rules about choosing one word, two-word or hyphenated writing. The boundaries of this class are vague: while some erratic spellings are obviously intentional and typical for Olbanian (say, writing a preposition and a noun as one word), some are very likely to be just unintentional errors (for instance, the rules which determine whether an adjective and a negative particle should be spelt as one word or as two words are extremely complex in Russian, and errors are often made in this position). Unfortunately, there is no way to distinguish between intentional and unintentional errors here, so all the erratic spellings of this kind were gathered under this type.

The "typo" type includes errors which seem to be unintentional, i.e. mostly random typos. To be considered a typo, a deviation must not belong to any other type, apart from "phonology". In theory, any phonological deviation is possible, so additional criteria are required to distinguish between typos and intentional phonological substitutions. The following was considered to be evidence of a typo: 1) the deviation occurs only once or twice; 2) it does not seem to belong to any established errativization pattern (i.e. hardening of soft consonants); 3) the keyboard layout suggests it is a probable typo.

Deviation from the rules of graphemics, i.e. using *3* instead of the Cyrillic letter *з / z*, did occur in the data, but were not included in the analysis due to their small number.

It is not quite clear whether the minor types (apart from "Olbanian morphology") can actually be considered as belonging to the Olbanian language. It is difficult to

determine how intentional these deviations are, and, even if they are intentional, was the intention of the speaker to show that they were using Olbanian? Or should any deviation from the norm be considered to be a manifestation of the Olbanian spirit? I do not attempt to provide exact answers to these questions, and make decisions only about the major types and "Olbanian morphology" (definitely Olbanian) and "typo" (definitely non-Olbanian). I view other types as containing both deviations that occurred as a result of Olbanian influence and without it, but the exact proportion remains unknown.

A3. Preved-hypothesis

Hristova (2011) insists on distinguishing two sociolects of Olbanian: the *padonki* language and the *preved*-phenomenon. The former emerged in the late 1990s and relies on the transcription class of the "orthography" type, while the latter emerged in 2006, triggered by a single preved-meme,[72] and follows the broader principle "write contrary to standard Russian", i.e. relies on the whole "orthography" type, including "anti-transcription" class.[73] She acknowledges that these two sociolects influence each other and the boundaries between them become blurred.

These claims can be verified by answering the following question: *Is it true that in 2006 the percentage of anti-transcription class increases sharply?*

[72] A painting by American artist John Lurie with the Russian word *preved* on it.

[73] Not all statements are explicit in Hristova's article, so all possible errors of interpretation are mine.

A4. Results

Relative frequencies (number of erratic forms/number of texts produced within the given year) for all the deviation types and all the years are summarized in Table A1. To provide a better illustration of the actual distributions, results for the three major types are presented in Fig. A1 as boxplots.

	2001	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Orthography: transcription	15.35	3.08	3.57	2.34	2.65	2.48	2.18	1.60	2.14	1.00	0.38
Orthography: anti-transcription	1.44	0.34	0.48	0.38	0.55	0.71	1.01	0.45	0.52	0.29	0.20
Phonology	0.56	0.17	0.19	0.17	0.24	0.28	0.18	0.31	0.16	0.11	0.11
One word or two words	0.38	0.56	0.59	0.29	0.73	0.58	0.50	0.44	0.32	0.38	0.38
Vernacular	0.18	0.02	0.07	0.12	0.02	0.08	0.02	0.02	0.11	0.00	0.02
Typo	0.32	0.19	0.27	0.12	0.08	0.02	0.09	0.02	0.04	0.11	0.05
Portmanteau	0.00	0.02	0.07	0.00	0.03	0.01	0.02	0.00	0.00	0.05	0.00
Common morphology	0.06	0.06	0.04	0.00	0.00	0.02	0.01	0.16	0.00	0.04	0.00
Olbanian morphology	0.00	0.00	0.00	0.01	0.01	0.09	0.11	0.03	0.00	0.00	0.00
Σ	18.29	4.44	5.28	3.43	4.31	4.27	4.12	3.03	3.29	1.98	1.14

Table A1. Frequency of deviations of different types (per 100 words)

It can be seen from Table A1 that major types (especially the “transcription” class) follow the pattern demonstrated by the total frequency in Fig. 1 in the main text: a

very sharp decrease after 2001 and a gradual decrease afterwards. A notable exception is comes by the “anti-transcription” class where a stable rebound is seen in the years 2005–2007. This can probably be attributed (at least for the years 2006 and 2007) to the influence of the preved-meme described by Hristova. Note, however, that one cannot claim that this deviation type *emerged* in 2006: it had existed previously, moreover, it was at its most popular in 2001.

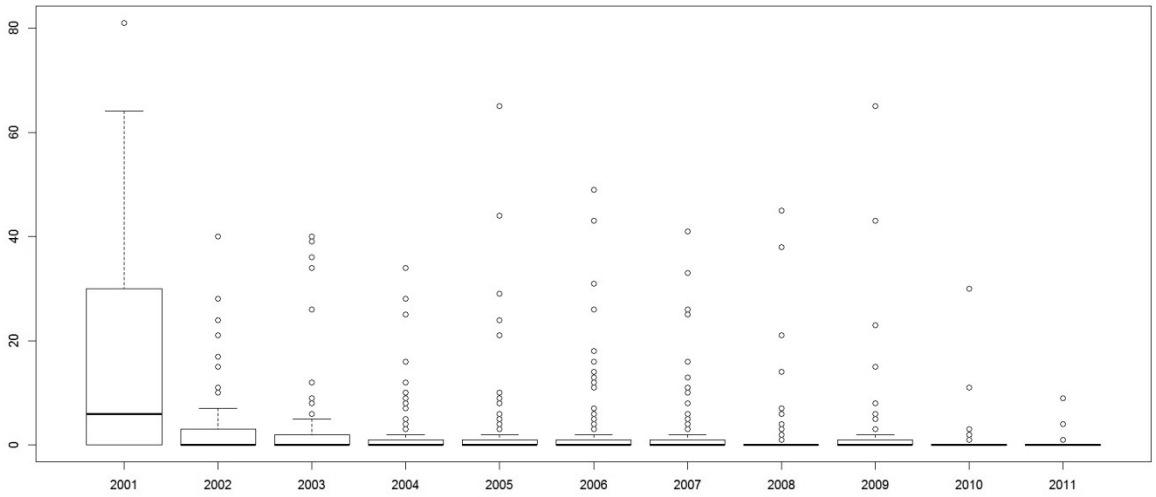

Figure A1a. Number of deviations (“orthography” type, “transcription” class) per 100 words

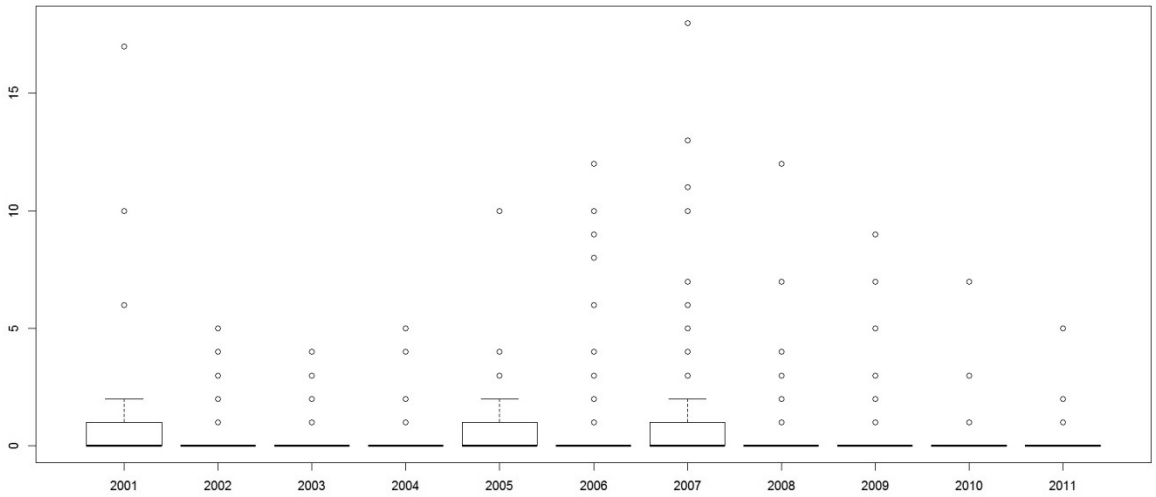

Figure A1b. Number of deviations (“orthography” type, “anti-transcription” class) per 100 words

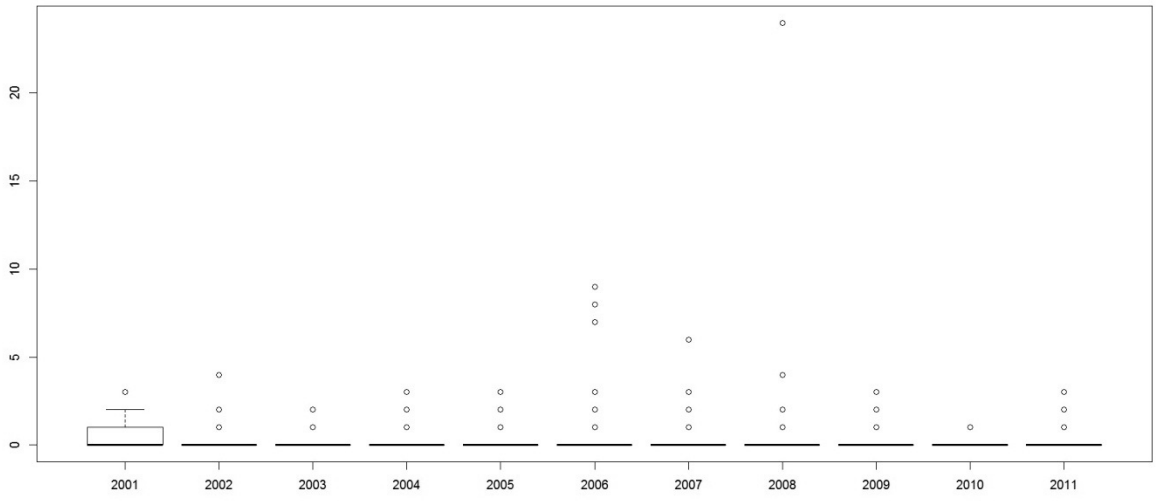

Figure A1c. Number of deviations ("phonology" type) per 100 words

The same process is probably responsible for the increase in the "Olbanian morphology" type in 2006 and 2007. The most popular morphological deviation during these two years is of the type described above: the "anti-transcription" spelling in nominative singular (*kotik* → *koteg*) causes an analogous alignment in other word forms (*kotika* → *kotega*). Thus, if the "anti-transcription" class gains in frequency and salience, this might easily stimulate the increase in the "Olbanian morphology" type.

The "one word or two words" is a vague case. The relative stability is probably ensured by a large share of unintentional errors, and it is reasonable to expect their rate to be constant, but I am not sure how to interpret the temporary increase in the middle.

Since the other types do not actually belong to the Olbanian language, we can try to use them for the purpose of comparison. The "typo" type demonstrates a general decrease. One possible explanation might be that writing in "proper" Olbanian, as in 2001, is actually quite difficult, and, when concentrating on this task, writers produce more typos. This offers some evidence to counter the popular assumption that

Olbanian is mostly used (or at least was used at the beginning) by highly literate people (Sokolovskij 2008, Hristova 2011) who allowed as many erratic spellings as they wanted, but committed no errors beyond that.

Other types do not demonstrate any clear patterns. The increase in "common morphology" in 2008 is caused by a single text and for that reason is not of much interest.

Finally, Fig. A2 demonstrates relative shares of the three major types and all the minor types combined in each year.

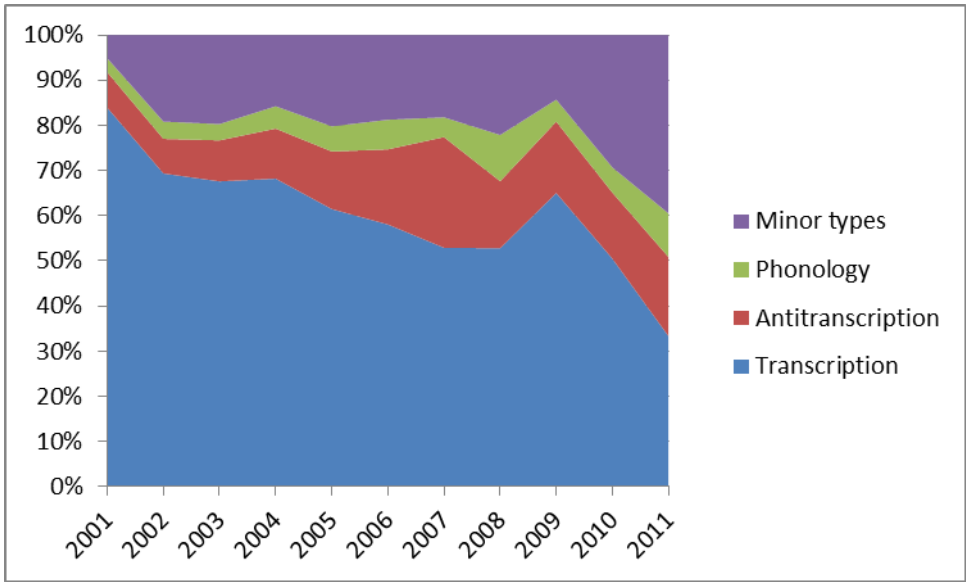

Figure A2. Relative contribution of different deviation types per year.

It is clear that the "transcription" class has always been dominant, but is gradually losing its importance. "Anti-transcription" and "phonology" are somewhat increasing their presence. Interestingly, "anti-transcription" reaches its maximum in 2006 and 2007, which again confirms the influence of the preved-meme. The contribution of the minor types, which include a large share of non-Olbanian deviations, increases greatly towards the end of the period.

A5. Conclusion

We can clearly see that the “transcription” class of the “orthography” type is responsible for the largest share of deviations. This is unsurprising, since *padonki* originally protested not against language, but against linguistic norm. Orthography, at least in the case of Russian, is the embodiment of the norm: an artificial and rigid system, codified more strictly than any other language domain and quite difficult to learn due to the lack of one-to-one mapping with oral speech.

The absolute decrease mostly affects the dominant “transcription” class, which leads to an increase in the shares of other types. The absolute and relative increase in the “anti-transcription” class in 2006–2007 is probably stimulated by the preved-meme, but the class is not created by it: it existed earlier. In actual fact, it is most likely that its previous existence ensured the high success of the meme: if it had been using a completely unfamiliar kind of deviation, it may have experiences difficulties achieving popularity.

References (combined list)

Aaron, P. and R. Malatesha Joshi. 2006. “Written language is as natural as spoken language: a biolinguistic perspective.” *Reading Psychology* 27: 263–311.

American Psychological Association. 2010. *Publication manual of the American Psychological Association, 6th edition*. Washington, DC: APA.

Andersen, Henning. 1989. “Understanding linguistic innovations”. In *Language change: contributions to the study of its causes*, edited by Leiv Egil Breivik and Ernst Håkon Jahr, 5–28. Berlin: Mouton de Gruyter.

Andersen, Henning, ed. 2001. *Actualization: linguistic change in progress*. Amsterdam, Philadelphia: Benjamins.

Andersen, Henning. 2006. “Synchrony, Diachrony and Evolution.” In *Competing Models of Linguistic Change: Evolution and Beyond*, edited by Ole Nedergaard Thomsen, 59–90. Amsterdam, Philadelphia: Benjamins.

Anderson, Jeffrey, Fred Beard and Joseph Walther. 2010. “Turn-Taking and the Local Management of Conversation in a Highly Simultaneous Computer-Mediated Communication System” *Language@Internet* 7, article 7. http://www.languageatinternet.org/articles/2010/2804, accessed January 8, 2013.

Androutsopoulos, Jannis. 2008. “Potentials and Limitations of Discourse-Centred Online Ethnography”. *Language@Internet 5*. Accessed November 27, 2012. http://www.languageatinternet.org/articles/2008/1610

Androutsopoulos, Jannis. 2011a. “Language change and digital media: A review of conceptions and evidence”. In *Standard languages and language standards in a changing Europe*, edited by Kristiansen Tore and Nikolas Coupland, 145–161. Oslo: Novus.

Androutsopoulos, Jannis. 2011b. “From variation to heteroglossia in the study of computer-mediated discourse.” In *Digital Discourse: Language in the New Media* edited by Crispin Thurlow and Kristine Mroczek, 277–298. Oxford: OUP.

Androutsopoulos, Jannis. 2006. “Introduction: Sociolinguistics and computer-mediated communication.” *Journal of Sociolinguistics*, 10(4): 419–438.

Anis, Jacques. 2007. “Neography: Unconventional spelling in French SMS text messages.” In *The Multilingual Internet. Language, Culture, and Communication Online*, edited by Brenda Danet and Susan Herring, 87–115. Oxford: OUP.

Anni, Oja. 2008. “Choosing language in Internet conversations between Russians and Estonians.” In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference “Dialogue”* 7 (14), 602–605. Moscow: RGGU.

Baron, Naomi. 1984. “Computer Mediated Communication as a Force in Language Change.” *Visible language* 18 (2): 118–141.

Baron, Naomi. 2004. “See You Online: Gender Issues in College Student Use of Instant Messaging.” *Journal of Language and Social Psychology* 23: 397–423.

Baron, Naomi. 2005. “The written turn.” Review of *The English writing system*, by Vivian Cook. *English Language and Linguistics* 9:359-376

Baron, Naomi. 2010. “Discourse Structures in Instant Messaging: The Case of Utterance Breaks.” *Language@Internet* 7, article 4. http://www.languageatinternet.org/articles/2010/2651, accessed January 8, 2013.

Baym, Nancy. 1995. “The Performance of Humor in Computer-Mediated Communication.” *Journal of Computer-Mediated Communication* 1 (2).

Bechar-Israeli, Haya. 1995. “From 〈Bonehead〉 TO 〈cLoNehEAd: nicknames, play, and identity on Internet Relay Chat”. *Journal of Computer-Mediated Communication* 1(2).

Beißwenger, Michael. 2008. “Situated Chat Analysis as a Window to the User’s Perspective: Aspects of Temporal and Sequential Organization.” *Language@Internet* 5, article 6. http://www.languageatinternet.org/articles/2008/1532, accessed January 23, 2013.

Belz, Julie and Jonathon Reinhardt. 2004). “Aspects of advanced foreign language proficiency: Internet-mediated German language play”. *International Journal of Applied Linguistics* 14(3): 324–362.

Berdičevskij, Aleksandr. 2012. “Âzykovaâ složnost′ [Language complexity. A review]. *Voprosy âzykoznaniâ* 5: 101–124.

Berglund, Therese. 2009. “Disrupted turn adjacency and coherence maintenance in instant messaging conversations”. *Language@Internet*, 6, article 2. http://www.languageatinternet.de/articles/2009/2106, last accessed January 8, 2013.

Biber Douglas. 1998. *Variation across speech and writing*. Cambridge: CUP.

Biber, Douglas. 1994. “An Analytical Framework for Register Studies”. In *Sociolinguistic perspectives on Register*, edited by Douglas Biber and Edward Finegan, 44–56. New York: OUP.

Blashki, Katherine, and Sophie Nichol. 2005. Game Geek’s Goss: Linguistic Creativity In Young Males Within An Online University Forum (94/\/\3 933k’5 9055oneone). *Australian Journal of Emerging Technologies and Society* 3(2): 77–86.

Bloomfield, Leonard. 1933. *Language* (revised from 1914 edition). New York: Holt.

Bogdanov, Aleksej. 2008. “Orfografiâ v Internete: analiz odnoj orfografičeskoj ošibki.” [Ortography at the Internet: the analysis of one orthographical error] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference “Dialogue”* 7 (14), 50–56. Moscow: RGGU.

Buras, Mariâ, and Maksim Krongauz. 2007. “Âzyk korporativnyh sajtov: igra, parodiâ, provokaciâ.” [The language of the corporate websites: play, parody,

provocation] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"*, 109–114. Moscow: RGGU.

Burkhaller, Byron, Marc Smith, and JJ Cadiz. 2000. Conversation trees and threaded chats. In *CSCW '00 Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 97–105. New York: ACM.

Bushman, Brad and L. Rowell Huesmann. 2012. "Effects of Televised Violence on Aggression." In *Handbook of Children and the Media*, edited by Dorothy Singer and Jerome Singer, 223–254. London: Sage.

Butcher, Andy. 2006. "Australian Aboriginal Languages: Consonant-Salient Phonologies and the 'Place-of-Articulation Imperative'." In *Speech Production: Models, Phonetic Processes, and Techniques*, edited by Jonathan Harrington and Marija Tabain, 187–210. New York: Psychology Press.

Chafe, Wallace. 1982. "Integration and Involvement in Speaking, Writing, and Oral Literature." In *Spoken and Written Language: Exploring Orality and Literacy*, edited by Deborah Tannen, 35–53. Norwood, NJ: Ablex.

Cherny, Lynn. 1999. *Conversation and community: Chat in a virtual world*. Stanford: Center for the Study of Language and Information.

Cheshire, Jenny, David Agder and Sue Fox. 2012. "Relative *who* and the actuation problem." *Lingua*, Available online 9 January 2013, doi 10.1016/j.lingua.2012.11.014.

Chomsky, Noam. 1959. "A review of B.F. Skinner's *Verbal Behavior*." *Language* 35 (1): 26–58.

Chovanec, Jan. 2011. "Humour in quasi-conversations: constructing fun in online sports journalism." In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 243–264. Amsterdam, Philadelphia: John Benjamins.

Collot, Milena and Nancy Belmore. 1996. “Electronic language: A new variety of English”. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 13–28. Amsterdam: Benjamins.

Condon, Sherri and Claude Čech. 2010. “Discourse Management in Three Modalities.” *Language@Internet* 7, article 6. http://www.languageatinternet.org/articles/2010/2770, accessed January 24, 2013.

Coulmas, Florian. 1989. *The Writing Systems of the World.* Oxford: Blackwell.

Croft, William. 2000. *Explaining language change: an evolutionary approach.* Harlow: Longman.

Croft, William. 2000. *Explaining language change: an evolutionary approach.* Harlow: Longman.

Crystal, David. 2006. *Language and the Internet* (2nd ed.). Cambridge: CUP.

Dahlberg, Lincoln. 2004. “Internet Research Tracings: Towards Non-Reductionist Methodology.” *Journal of Computer-Mediated Communication* 9(3) http://jcmc.indiana.edu/vol9/issue3/dahlberg.html

Danet, B., Ruedenberg-Wright, L. and Rosenbaum-Tamari, Y. (1997). “«HMMM…WHERE'S THAT SMOKE COMING FROM?». Writing, Play and Performance on Internet Relay Chat”. *Journal of Computer-Mediated Communication* 2 (4).

Danet, Brenda and Susan Herring. 2007. *The Multilingual Internet*. Oxford: OUP.

Danet, Brenda. 2001. *Cyberpl@y: Communicating Online*. Oxford: Berg.

Deuel, Nancy. 1996. “Our passionate response to virtual reality.” In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 129–146. Amsterdam, Philadelphia: John Benjamins.

Deutscher, Guy. 2000. *Syntactic change in Akkadian: the evolution of sentential complementation*. Oxford: OUP.

Dresner, Eli. 2005. "The Topology of Auditory and Visual Perception, Linguistic Communication, and Interactive Written Discourse." *Language@Internet* 2, article 2. http://www.languageatinternet.org/articles/2005/161, accessed January 23, 2013.

Dynel, Marta. 2009. "Beyond a joke: types of conversational humour." *Language and Linguistics Compass* 3 (5): 1284–1299.

Dynel, Marta. 2011a. "Pragmatics and linguistic research into humour." In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 1–15. Amsterdam, Philadelphia: John Benjamins.

Dynel, Marta. 2011b. "Joker in the pack: towards determining the status of humorous framing in conversations." In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 271–242. Amsterdam, Philadelphia: John Benjamins.

Eble, Connie. 1996. *Slang and sociability: in-group language among college students.* Chapel Hill: UNC Press.

Fahlman, Scott. ?2002. "Smiley lore." http://www.cs.cmu.edu/smiley/history.html, accessed 20 January 2013.

Ferrara, Kathleen, Hans Brunner and Greg Whittemore. 1991. "Interactive written discourse as an emergent register." *Written Communication* 8 (1): 8–34.

Fought, John, Robert Munroe, Carmen Fought, and Erin Good. 2004. "Sonority and Climate in a World Sample of Languages: Findings and Prospects." *Cross-Cultural Research* 38 (1): 27–51.

Fung, Loretta and Ronald Carter. 2007. "Cantonese e-discourse: A new hybrid variety of English." *Multilingua* 26: 35–66.

Gao, Liwei. 2006. “Language contact and convergence in computer-mediated communication.” *World Englishes* 25 (2): 299–308.

Garcia, Angela, and Jennifer Jacobs. 1999. “The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication.” *Research on Language & Social Interaction* 32 (4): 337–367.

Gee, James. 1986. “Orality and Literacy: From *The Savage Mind* to *Ways With Words.*” *TESOL Quarterly* 20 (4): 719–746.

Georgakopolou, Alexandra. 2006. “Postscript: Computer-mediated communication in sociolinguistics”. *Journal of Sociolinguistics* 10 (4): 548–557

Gibson, James. 1977. “The theory of affordances.” In *Perceiving, acting, and knowing*, edited by R. E. Shaw & J. Brandsford, 67–82. Hillsdale, NJ: Lawrence Erlbaum.

González-Lloret, Marta. 2011. “Conversation Analysis of Computer-Mediated Communication.” *CALICO Journal* 28 (2): 308–325.

Goodwin, Charles and John Heritage. 1990. “Conversation Analysis.” *Annual Review of Anthropology* 19: 283–307.

Gottlieb, Nanette. 2010. “Playing with Language in E-Japan: Old Wine in New Bottles.” *Japanese Studies* 30 (3): 393–407.

Greiffenstern, Sandra. 2009. The Influence of Computers, the Internet and Computer-Mediated Communication on Everyday English. Berlin: Logos Verlag.

Gusejnov, Gasan. 2005. “Berloga vebloga: Vvedenie v èrratičeskuû semantiku.” [The lair of the weblog: Introduction into the erratic semantics] http://www.speakrus.ru/gg/microprosa_erratica-1.htm

Halliday, M.A.K. 1976. “Anti-Languages.” *American Anthropologist* 78(3): 570–584.

Hård af Segerstad, Ylva. 2002. “Use and Adaptation of Written Language to the Conditions of Computer-Mediated Communication.” PhD diss., University of Gothenburg.

Haspelmath, Martin. 2010. “Framework-free grammatical theory.” In *The Oxford handbook of grammatical analysis*, edited by Bernd Heine and Heiko Narrog, 341–365. Oxford: OUP.

Herring, Susan. 1998. “Le style du courrier électronique: variabilité et changement.” *Terminogramme* 84-85: 9-16. English version accessed 23 January 2013 at http://ella.slis.indiana.edu/~herring/term.html.

Herring, Susan. 1999. “Interactional coherence in CMC.” *Journal of Computer-Mediated Communication* 4 (4).

Herring, Susan. 2003. “Media and language change: introduction.” *Journal of Historical Pragmatics* 4 (1): 1–17.

Herring, Susan. 2007. “A Faceted Classification Scheme for Computer-Mediated Discourse.” *Language@Internet*, 4, article 1.
http://www.languageatinternet.org/articles/2007/761, accessed January 8, 2013.

Herring, Susan. 2012. “Grammar and electronic communication.” In *The Encyclopedia of Applied Linguistics*, edited by Carol Chapelle. Hoboken, New Jersey: Wiley-Blackwell.
http://onlinelibrary.wiley.com/book/10.1002/9781405198431

Hristova, Daniela. 2011. “Velikij i mogučij olbanskij âzyk: The Russian Internet and the Russian Language.” *Russian Language Journal* 143–162.

Hutchby, Ian. 2001. *Conversation and Technology: From the Telephone to the Internet*. Cambridge: Polity.

Il′ina, N. E. 1996. “Rost analitizma v morfologii.” [Analytic trends in Russian morphology] In *Russkij âzyk konca XX stoletiâ (1985–1995)* [Russian language at the

end of 20th century (1985-1995)], edited by Elena Zemskaâ, 326–343. Moscow: Âzyki russkoj kul′tury.

Itkin, Il′a. 2007. *Russkaâ morfonologiâ* [Russian morphonology] Moscow: Gnozis.

Jarbou, Samir and Buthaina al-Share. 2012. "The Effect of Dialect and Gender on the Representation of Consonants in Jordanian Chat." *Language@Internet* 9, article 1. http://www.languageatinternet.org/articles/2012/Jarbou, accessed January 24, 2013.

Jensen, Bård Uri. 2008. "Syntactic variables in pupils' writing: A comparison between hand-written and pc-written texts." In *Linguistics in the making: selected papers from the second Scandinavian PhD conference in linguistics and philology in Bergen, June 4-6, 2007*, edited by Gard B. Jenset, Øystein Heggelund, Margrete Dyvik Cardona, Stephanie Wold, and Anders Didriksen, 165–184. Oslo: Novus Press.

Jones, Graham and Schieffelin, Bambi. 2009. "Enquoting voices, accomplishing talk: Uses of *be* + *like* in instant messaging." *Language & Communication* 29: 77–113.

Karlsson, Fred. 2009. "Origin and maintenance of clausal embedding complexity." In *Language complexity as an evolving variable*, edited by David Gil, Geoffrey Sampson and Peter Trudgill. Oxford: OUP.

Keating, Elizabeth and Gene Mirus. 2003. "American Sign Language in virtual space: Interactions between deaf users of computer-mediated video communication and the impact of technology on language practices." *Language in Society* 32: 693–714.

Kibrik, Aleksandr. 1992. *Očerki po obŝim i prikladnym voprosam âzykoznaniâ* [Essays on general and applied aspects of linguistics] Moscow: MGU.

Knâzev, Sergej, and Sof̲ʹâ Požarickaâ. 2012. *Sovremennyj russkij literaturnyj âzyk. Fonetika, orfoepiâ, grafika i orfografiâ.* [Contemporary standard Russian language.

Phonetics and phonology, orthoepy, graphics and orthography] Moscow: Gaudeamus.

Ko, Kwang-Kyu. 1996. “Structural Characteristics of Computer-Mediated Language: A Comparative Analysis of InterChange Discourse.” *Electronic Journal of Communication* 6 (3)

Krongauz, Maksim. 2008. “Utomlennye gramotoj.” [Wearied by literacy] *Novyj mir* 5: 153–159. Accessed at http://magazines.russ.ru/novyi_mi/2008/5/kr11.html

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

Labov, William. 1994. *Principles of linguistic change, Internal factors*. Hoboken: John Wiley & sons.

Labov, William. 2010. *Principles of Linguistic Change, Cognitive and Cultural Factors*. Hoboken: John Wiley & sons.

Lancaster, Sean, David C. Yen, Albert H. Huang, and Shin-Yuan Hung. 2007. “The selection of instant messaging or e-mail: College students’ perspective for computer communication”. *Information Management & Computer Security* 15(1): 5–22.

LaPolla, Randy. 2003. “Why languages differ: Variation in the conventionalisation of constraints on inference.” In *Language variation: Papers on variation and change in the Sinosphere and in the Indosphere in honour of James A. Matisoff*, edited by D. Bradly, R. LaPolla, B. Michailovsky, G. Thurgood, 113–144. Canberra: Pacific Linguistics.

Lâševskaâ, Olga; and Šarov, Sergej. 2009. *Novyj častotnyj slovar' russkoj leksiki* [New frequency dictionary of Russian words]. Moscow: Azbukovnik. http://dict.ruslang.ru/freq.php

Lee, Carmen K.-M. 2007.” Affordances and Text-Making Practices in Online Instant Messaging.” *Written Communication* 24: 223–249.

Levin, Maurice. 1978. *Russian declension and conjugation: a structural description with exercises*. Columbus: Slavica Publishers.

Liberman, Mark. 2011. "Telegraphic Language" *The Language Log*, December 31. http://languagelog.ldc.upenn.edu/nll/?p=3674

Lighter, Jonathan. 1994. Preface to *Historical Dictionary of American Slang*, xi–xxxix. Oxford: OUP.

Lillis, Theresa. 2013. *The Sociolinguistics of Writing.* Edinburgh: Edinburgh University Press.

Ling, Richard, and Naomi Baron. 2007. "Text Messaging and IM: Linguistic Comparison of American College Data." *Journal of Language and Social Psychology* 26 (3): 291–299.

Lûdovik, T. 2010. Analiz tekstov SMS-soobŝenij s celʹû povyšeniâ kačestva [Analysis of SMS texts aimed at improving their automatical reading] // In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"* 9 *(*16), 313–317. Moscow: RGGU.

Makarov, M., and M. Školovaâ. 2006. "Lingvističeskie i semiotičeskie aspekty konstruirovaniâ identičnosti kommunikacii." [Linguistic and semiotic aspects of identity construction in e-communication] *Computational linguistics and intellectual technologies*. *Papers from the annual international conference "Dialogue"*, 364–369. Moscow: RGGU.

Marcoccia, Michel, Hassan Atifi and Nadia Gauducheau. 2008. "Text-Centered versus Multimodal Analysis of Instant Messaging Conversation." *Language@Internet* 5, article 7.
http://www.languageatinternet.org/articles/2008/1621, accessed January 24, 2013.

Markov, Aleksandr. 2010. *Rožden*ie *složnosti*. *Èvolûcionnaâ biologiâ segodnâ: neožidannye voprosy i novye otkrytiâ* [The emergence of complexity. Evolutionary biology today: unexpected questions and new discoveries]. Moscow: Astrel′.

Maslova, Elena. 2008. "Unidirectionality of grammaticalization in an evolutionary perspective." In *Studies on grammaticalization*, edited by Elisabeth Verhoeven, Stavros Skopeteas, Yong-Min Shin, Yoko Nishina, Johannes Helmbrecht, 15–24. Berlin, New York: Mouton de Gruyter.

Matras, Yaron. 2009. *Language Contact*. Cambridge: CUP.

Mechkovskaia, Nina. 2006. "Estestvennyk iazyk i metaiazykovaâ refleksiâ v vek Interneta [Natural language and metalinguistic reflection in the era of the Internet]." *Russkij âzyk v nauchnom osveŝenii*, 12 (2): 165–185.

Milroy, James, and Lesley Milroy. 1985. "Linguistic change, social network and speaker innovation." *Journal of Linguistics* 21 (2): 339–384.

Morkes, John, Hadyn Kernal, and Clifford Nass. 1999. "Effects of Humor in Task-Oriented Human-Computer Interaction and Computer-Mediated Communication: A Direct Test of SRCT Theory" *Human-Computer Interaction* 14 (4): 395–435.

Murray, Denise. 1991. *Conversation for action*: *The computer terminal as medium of communication*. Amsterdam, Philadelphia: Benjamins.

Naper, Ida. 2011. "Conversation in a multimodal 3D virtual environment." *Language@Internet* 8, article 7. http://www.languageatinternet.org/articles/2011/Naper, Accesses January 24, 2013.

Nejdaeva, Svetlana. 1985. "Istoriâ grammatičeskogo razvitiâ suŝestvitel′nyh sub″ektivnoj ocenki." [The history of the grammatical development of the subjective evaluation nouns] PhD diss., Alma-Ata.

Norman, Boris. 2006. “Lingvopsihologičeskie aspekty grammatičeskoj kategorii roda.” [Psycholinguistic facets of the grammatical category of gender] Russian Linguistics 30(2): 153–174.

Norman, Don. 1990. *The design of everyday things*. New York: Doubleday.

Norman, Don. 1999. Affordance, conventions and design. *Interactions* (May/June): 39-42.

Norrick, Neal. 2010. “Humor in interaction.” *Language and Linguistics Compass* 4 (4): 232–244.

Öberg, Mattias, Maritta S Jaakkola, Alistair Woodward, Armando Peruga, and Annette Prüss-Ustün. 2011. “Worldwide burden of disease from exposure to second-hand smoke: a retrospective analysis of data from 192 countries.” *The Lancet* 377 (9760): 139–146.

Olson, David. 1996. “Towards a psychology of literacy: on the relations between speech and writing.” *Cognition* 60: 83–104.

Palfreyman, David, and al Khalil, Muhamed. 2007. “‘A Funky Language for Teenzz to Use’: Representing Gulf Arabic in Instant Messaging.” In *The Multilingual Internet*, edited by Brenda Danet and Susan Herring, 43–63. Oxford: OUP.

Paolillo, John. 2001. “Language variation on Internet Relay Chat: A social network approach.” *Journal of Sociolinguistics* 5 (2): 180–213.

Paulsen, Martin. 2011. “Digital determinism: the Cyrillic Alphabet in the Age of New Technology.” *Russian Language Journal* 61: 119–142.

Perea, Manuel, Jon Andoni Duñabeitia, and Manuel Carreiras. 2008. “OBSERVATIONS: R34D1NG W0RD5 W1TH NUMB3R5.” *Journal of Experimental Psychology Human Perception and Performance* 34 (1): 237–241.

Perry, Fred. 2005. *Research in applied linguistics: becoming a discerning consumer*. Mahwah: Laurence Erlbaum Associates.

Plester, Beverly, and Clare Wood. 2009. "Exploring Relationships Between Traditional and New Media Literacies: British Preteen Texters at School." *Journal of Computer-Mediated Communication* 14(4): 1108–1129.

Polinsky, Maria. 2006. "Incomplete Acquisition: American Russian." *Journal of Slavic Linguistics* 14 (2): 191–262.

Popescu, Carmen. 2011. "Understanding ethnic humour in Romanian jokes". In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 173–190. Amsterdam, Philadelphia: John Benjamins.

Rouzie, Albert. 2001. "Conversation and Carrying-On: Play, Conflict, and Serio-Ludic Discourse in Synchronous Computer Conferencing." *College Composition and Communication* 53 (2): 251–299.

Rowe, Charley. 2011. "*Whatchanade?* Rapid language change in a private email sibling code." *Language@Internet* 8, article 6. http://www.languageatinternet.org/articles/2011/Rowe, accessed January 23, 2013.

Sacks, Harvey. 1974. "An analysis of the course of a joke's telling in conversation." In *Explorations in the Ethnography of Speaking*, 337–353, edited by Richard Bauman and Joel Sherzer. Cambridge: CUP.

Sandler, Wendy, and Lillo-Martin, Diane. 2006. *Sign language and linguistic universals.* Cambridge: CUP.

Sannikov, Vladimir. 2002. *Russkij âzyk v zerkale âzykovoj igry* [Russian language in the mirror of language play] (2nd ed.). Moscow: Âzyki slavânskoj kul′tury.

Šapovalova, Nadežda. 2007. "Karnaval′noe obŝenie v Internete (na materiale sajta www.udaff.com)." [Carnivalesque communication in the Internet (using the data of

the website www.udaff.com] *Vestnik Čelâbinskogo gosudarstvennogo universiteta* 16 (20): 164–167.

Saussure, Ferdinand. 2011. *Course in general linguistics*. Translated by Wade Baskin, edited by Perry Meisel and Haun Saussy. New York: Columbia University Press.

Sebba, Mark. 2003. "Spelling rebellion." In *Discourse constructions of youth identities*, edited by Jannis Androutsopoulos and Alexandra Georgakopoulou, 151–172. Amsterdam, Philadelphia: Benjamins.

Sebba, Mark. 2009. Sociolinguistic approaches to writing systems research. *Writing Systems Research* 1 (1): 35–49.

Sidorova, Marina. 2006. *Internet-lingvistika: russkij âzyk. Mežličnostnoe obŝenie.* [Internet-linguistics: Russian language. Interpersonal communication]. Moscow: 1989.ru

Simpson, James. 2005. "Conversational floors in synchronous text-based CMC discourse." *Discourse Studies* 7(3): 337-361.

Skinner, B.F. 1957. *Verbal Behavior.* New York: Appleton-Century-Crofts.

Smith, Bryan and Greta Gorsuch. 2004. "Synchronous computer mediated communication captured by usability lab technologies: New interpretations." *System* 32: 553-575.

Smyk-Bhattacharjee, Dorota. 2006. "'A 'lyricist' by night, teenager by day and sleepaholic everytime in between'. Language change in/through CMC". *Germanistische Linguistik* (186–187): 68–85.

Sokolovskij, Dmitrij (2008). *Bibliâ padonkov, ili Učebneg albanskogo âzyka*. Moscow: Folio SP.

Sparrow, Betsy, Jenny Liu, and Daniel M. Wegner. 2011. “Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips.” *Science* 333 (6043): 776–768.

Squires, Laurel. 2007. “Whats the use of apostrophes? Gender differences and linguistic variation in instant messaging.” *TESOL Working Papers* 4. Accessed January 24, 2013 at http://aladinrc.wrlc.org/bitstream/handle/1961/5240 /CMCSquiresFinal.pdf?sequence=1

Squires, Laurel. 2010. “Enregistering internet language.” *Language in Society* 39: 457–492.

Su, Hsi-Yao. 2003. “The Multilingual and Multi-Orthographic Taiwan-Based Internet: Creative Uses of Writing Systems on College-Affiliated BBSs.” *Journal of Computer-Mediated Communication* 9(1).

Tagliamonte, Sali, and Derek Denis. 2008. “Linguistic ruin? LOL! Instant messaging and teen language.” *American Speech* 83(1): 3–34.

Tagliamonte, Sali. 2006. *Analysing sociolinguistic variation*. Cambridge: CUP.

ten Have, Paul. 1986. “Methodological issues in conversation analysis.” Bulletin de Méthodologie Sociologique, 27: 23–51. Accessed at http://www.paultenhave.nl/mica.htm#N_1_, January 8, 2013.

Testelec, Âkov. 2001. *Vvedenie v obŝij sintaksis* [An introduction to general syntax] Moscow: RGGU.

The “Five Graces Group” (Clay Beckner, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann). 2009. “Language Is a Complex Adaptive System: Position Paper”. *Language Learning* 59, s1: 1–26.

Thielemann, Nadine. 2011. “Displays of “new” gender arrangements in Russian jokes.” In *The pragmatics of humour across discourse domains*, edited by Marta Dynel, 147–172. Amsterdam, Philadelphia: John Benjamins.

Thurlow, Crispin. 2006. “From statistical panic to moral panic: The metadiscursive construction and popular exaggeration of new media language in the print media.” *Journal of Computer-Mediated Communication* 11:3. http://jcmc.indiana.edu/vol11/issue3/thurlow.html

Thurlow, Crispin. 2007. Fabricating youth: new-media discourse and the technologization of young people. *Language in the Media*, edited by S. Johnson and A. Ensslin, 213–233. London: Continuum.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: OUP

Vandergriff, Ilona, and Carolin Fuchs. 2009. “Does CMC Promote Language Play? Exploring Humor in Two Modalities.” *CALICO Journal* 27(1).

Varnhagen, Connie, G. Peggy McFall, Nicole Pugh, Lisa Routledge, Heather Sumida-MacDonald, and Trudy E. Kwong. 2010. "lol: new language and spelling in instant messaging" *Reading and Writing* 23 (6): 719–733.

Viktorova, M. 2008. “Èmotivnyj aspekt prodleniâ soglasnyh zvukov (na materiale pesen V. Vysockogo i Ž. Brelâ.” [The affective aspect of elongating the consonant sounds (on the data from the songs of Vladimir Vysotsky and Jacques Brel] *Izvestiâ Rossijskogo gosudarstvennogo pedagogičeskogo universiteta imeni A. I.* Gercena 60: 68–73.

Warner, Chantelle. 2004. “It’s just a game, right? Types of play in foreign language CMC.” *Language Learning and Technology* 8 (2): 69-87.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. *Empirical foundations for a theory of language change.* Austin: University of Texas Press.

Werry, Christopher. 1996. “Linguistic and interactional features of Internet Relay Chat. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 47–61. Amsterdam: Benjamins.

Wolfram, Walt. 1991. “The linguistic variable: fact and fantasy.” *American Speech* 66 (1): 22–32.

Wray, Alison and George Grace. 2007. “The consequences of talking to strangers: evolutionary corollaries of socio-linguistic influences on linguistic form.” *Lingua* 117 (3): 543–578.

Yang, Chunsheng. 2007. “Chinese Internet language: A sociolinguistic analysis of adaptations of the Chinese writing system.” *Language@Internet* 4, article 2. http://www.languageatinternet.org/articles/2007/1142. Accessed January 24, 2013.

Yates, JoAnne, and Wanda Orlikowski. 1993. “Knee-jerk Anti-LOOPism and other E-mail Phenomena: Oral, Written, and Electronic Patterns in Computer-Mediated Communication.” MIT Sloan School Working Paper, http://ccs.mit.edu/papers/CCSWP150.html

Yates, Simeon. 1996. “Oral and Written Linguistic Aspects of Computer Conferencing: A Corpus Based Study” In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by Susan Herring, 29–46. Amsterdam: Benjamins.

Zaliznâk, Andrej. 2004. *Drevnenovgorodskij dialekt*. [Old Novgorod dialect] Moscow: Âzyki slavânskoj kul′tury.

Zaliznâk, Anna, and Irina Mikaèlân. 2006. “Perepiska po èlektronnoj počte kak lingvističeskij ob″ekt.” [E-mail correspondence as an object of linguistic analysis] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference “Dialogue”*, 157–162. Moscow: RGGU.

Zaliznâk, Andrej. *Grammatičeskij slovar' russkogo âzyka. Slovoizmenenie* [Grammar dictionary of the Russian language. Inflection]. M., 1977.

Zanegina, N. 2009. “~~Â ètogo ne govoril~~: o liturarivah, začerkivaniâh ili mnimyh tekstah.” [~~I’ve never told that~~: about lituratives, strikeouts or imaginary texts] In *Computational linguistics and intellectual technologies*. *Papers from the annual international conference “Dialogue”* 8 (15), 112–115. Moscow: RGGU.

Zimmer, Ben. 2007. “The prehistory of emoticons.” *The Language Log*, September 21, 2007. http://itre.cis.upenn.edu/~myl/languagelog/archives/004935.html

Zitzen, Michaela, and Dieter Stein. 2004. “Chat and conversation: a case of transmedial stability?” *Linguistics* 42 (5): 983–1021.

Zubova, Lûdmila. 2010. “Ironičeskaâ grammatika: srednij rod v igrovoj neologii.” [Ironic grammar: neuter gender in playful neology] *Voprosy âzykoznaniâ* 6: 16–25.

Zubova, Lûdmila. 2000. “Kategoriâ roda i lingvističeskij èksperiment v sovremennoj russkoj poèzii.” [The gender category and the language experiment in modern Russian poetry] In *Problemy funkcional'noj grammatiki: Kategorii morfologii i sintaksisa v vyskazyvanii* [Problems of functional grammar: morphological and syntactic categories in the utterance], edited by A. Bondarko and S. Šubik, 194–210. St Petersburg: Nauka. Accessed at http://www.levin.rinet.ru/ABOUT/zubova1.html, January 26, 2013.

Zvereva, Vera. 2009. “«Âzyk padonkaf»: Diskussii pol′zovatelej Runeta.” [The *padonki* language: discussions of Runet users] In *From Poets to Padonki. Linguistic Authority & Norm Negotiation in Modern Russian Culture*, edited by Ingunn Lunde and Martin Paulsen, 49–79. Bergen.