Standards for measuring the Netspeaks quantity in on-line text content

Karmela, ALEKSIC-MASLAC
Zagreb School of Economics and Management
Zagreb, Croatia

Jagoda, POROPAT DARRER
Zagreb School of Economics and Management
Zagreb, Croatia

Tihana, DJURAS
Zagreb School of Economics and Management
Zagreb, Croatia

ABSTRACT

Everyday use of new communication channels such as MSN, Skype, Facebook, and SMS is changing the written language in many ways creating the new language form called Netspeak.

In this paper the authors develop the methodology for measuring the frequency in using Netspeak elements through ten specific standards. The standards are named, described and grouped into four categories regarding the provenience of language phenomena found. The first group of standards is related to the information communication technology; the second group is related to the grammar and syntax; the third one is related to the prosody and the fourth one named other is related to every other kind of provenience. The standards are applied in analyzing the content of asynchronous discussions throughout four generations of students within the course Information and Communication Technologies at Zagreb School of Economics and Management. The study shows the correlation between the use of standards within each group and shows the more frequent use of Netspeak elements by more active students.

Keywords – Netspeak elements, standards, on-line text content, discussion, quality

1. INTRODUCTION

"At the time Albert Einstein discovered the theory of relativity he faced a new surprising thing. This world known physicist soon realized that his fascinating and revolutionary discovery that shook the basics of Newton classic physics impossible to be expressed and explained using the existing vocabulary because of the very simple reason: the terms describing the new phenomenon didn't exist. Following the development of communication technologies and the transformation of media from auditory and written to the graphical and interactive, opened up the need for a specific language with a vocabulary that enables description and convey all the changes and phenomena in the communications field of today", says Ivana Tarnaj [1].

The authors of the European Council study "The new space of communication, the interface with culture and artistic activities" divide communications into two categories - interactive and mass communication. Interactive communication is based on the concept of shared space. When we communicate we share the same physical space with someone and we create an interactive system with a person with whom we are communicating and by using a common communication method which is the same language.

As well as technology, the notion of communication has a very tight connection to the notion of culture. In his book "Communication as Culture: Essays on Media and Society", James Carey explains the notion of communication through two models. The first model he names transmission, and the second one communication as a ritual. Carey describes communication as a transfer of concepts such as sending, transmission, providing information to others, transportation, etc., stating that communication is transmission of the signals or messages in distance for the purpose of control. The model of communication as ritual on the contrary refers not to dissemination of messages through space, but to the maintenance of society or community through time where communications is represented as a common belief. The communication as a ritual is a kind of a social ritual in which participants feel connected to each other in a community. [2]

In a study dealing with literacy in the computer age, Myron Tuman considers associability and nonlinearity as an attack on "the status of texts as a higher and more logical phrase of symbolic knowledge". Digital literacy is particularly problematic in education, which now encourages students to "cruise the information highway seeking and embracing the minimum necessary information", confusing information with knowledge. [3]

Fluid and volatile electronic environment, associative, non-hierarchical and non-linear organization is considered to be a benefit of the digital environment by other theorists. Multilinearity and dispersion are characteristics of the human psyche, which is why the new media fail "to imitate"

the human mind. The development of cyber-space, claim the advocates of technology, is an illustration of non-hierarchy and nonlinearity of the human perception. [4]

As shown in Figure 1, the goal of every communication process is to analyze the source, coder, transmitter, channel, receiver, decoder, and recipient. The communication process is set so that the source products the information (I) which is encoded in message (M). The transmitter materializes the message in the signal (S). Signal is good if it has the same shape as the message, if it is aligned with the channel, which also has very specific material properties. The signal in the channel is affected by noise (N), which interferes with communication flow. Noise or interference should be considered besides any discrepancy between the parts of the communication chain. Mismatch between the encoder and decoder generate semantic noise. The receiver can inform the source of the feedback notification (F) which closes the communication process. [5]

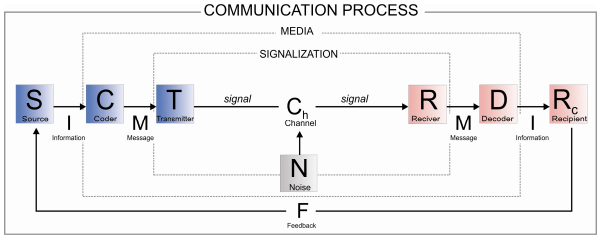

FIGURE 1
COMMUNICATION PROCESS

2. STANDARD DESCRIPTION

Croatian language is a Slavic language spoken by Croats in their communities. Any society that homogenizes, achieves it due to internal communication. For the purposes of this communication forms, a common language (koine) is created, which occurs spontaneously, appropriate to the needs and communications, space and/or time. It optimizes according to the principle of the minimum cost, the economy of the language, such as lingua franca. The standard idiom has a function of understanding and creating the official, general and cultural activities in schools, universities, parliament, national television and radio, print, public signs, services and activities of clerks and political, entrepreneurial, scientific and professional public. It is one and unique to all members of diverse organic idiom to whom is possible and desirable. [6]

In order to measure the quantity of Netspeak elements the authors created 10 standards grouped into 4 categories regarding their provenience. The first group is ICT and gathers 3 standards: words in English (I1), acronyms and abbreviations (I2), emoticons (I3). The second group is grammar and syntax and gathers also 3 standards: lower case graphemes (G1), diacritics (G2), space (G3). The third group is prosody and gathers 3 standards too: punctuation (P1), uppercase graphemes (P2), prolonged graphemes. The fourth group is „other" where the authors placed individual and sporadic elements such as the use of past tense „aorist", etc.

TABLE 1
STANDARDS FOR MEASURE THE QUALITY OF NETSPEAK ELEMENTS

STANDARD	DESCRIPTION	P
I1 – English words	New technologies development is based on English language so it happens that Croatian is subjected to overwhelming English words.	10
I2 – acronyms and abbreviations	Acronyms and abbreviations are composed of the initial letters of each member of the expression in them. Abbreviations are mixed; there are regular and occasional ones. There are common abbreviations that are short parts of words or sets of words, and read as if words are spelled correctly. Other abbreviations are formed by merging the initial letter or letters of multi-member group called names and is usually read as written.	10
I3 – emoticon	Emoticons are signs, symbols. They are not just colon and parentheses, it is a sign of a good or mood, and sometimes takes other meanings depending on the context in which it is used. Symbols are signs in which the relationship between signifiers are already learned.	10
G1 – lower case graphemes	Contrary to the grammar rules, the use of lower case graphemes where it should be used upper case graphemes.	10
G2 – diacritics special signs	Part of the grapheme that change the sound of the grapheme. Those signs are omitted and often recorded by the standard rules of English language.	10
G3 – space	The omission of space where needed, after punctuation.	10
P1 – punctuation	Punctuation is used in a non standard way in order to compensate the auditive channel within the discussion.	10
P2 –uppercase graphemes	In written Croatian language there is standard use of uppercase in three particular situations. First is with the proper names, the second as the first letter in a sentence and finally in order to express politeness. Though, there are some exceptions. Uppercase within the whole word, sentence or text can be used for esthetic, advertising or propaganda reasons. It is used in order to emphasize the specific word and to plan and to add the prosodic elements to the written word.	10
P3 –prolongation of the graphemes	In written Croatian language there are 30 sounds each represented by single grapheme (except three sounds being represented by double graphemes *dž, lj* and *nj*). There's no such a thing as geminate (a double consonant such as *mm* i.e. in word *communication*). It is used in order to add prosodic elements to written words. Prosody gives rhythm and melody to a word. It comprehends acoustic parameters such as accent, intonation, and melody.	10
O – Other	Use of tense considered to be obsolete – aorist. As far as the past tenses are concerned, the most frequent and the most dominant tense in contemporary Croatian is the Croatian *perfect - Vidjela sam te* (PERFECT – *to see*). Shortened form, *aorist* form would be *Vidjeh te.* (AORIST – *to see*).	10

3. MEASUREMENT OF THE QUANTITY OF NETSPEAK ELEMENTS WITHIN THE CLOSED DISCUSSION

The quantity determined under Netspeak ten standards will be measured at a very advanced closed discussion [7, 8] of the course Information and Communication Technologies [9] through 4 different generations. From 2008/2009 to 2011/2012, 421 students (39.64% of all students) have taken an active part in discussions. Analyzed sample is shown in Table 2.

TABLE 2
NUMBER OF STUDENTS WHO PARTICIPATED IN THE DISCUSSIONS

Academic year	∑ students participated in the discussion	∑ students on the course	%
2008/2009	134	295	45.42%
2009/2010	110	341	32.26%
2010/2011	103	244	42.21%
2011/20012	74	182	40.66%
∑	421	1062	39.64%

The most active generation of students is the one of the academic year 2008/09, when 45.42% of students actively participated in discussions as a supplementary activity within the course.

Table 3 shows the distribution of Netspeak elements grouped in ten standards through four generations of students analyzed within the same course.

TABLE 3
DISTRIBUTION OF NETSPEAK ELEMENTS IN TEN STANDARDS

	2008/ 2009	**2009/ 2010**	**2010/ 2011**	**2011/ 2012**	avg.
I1	92.62	83.65	94.53	96.84	95.50
I2	66.28	58.83	64.27	66.22	65.09
I3	26.80	32.33	28.14	36.16	31.49
G1	19.89	28.20	17.59	15.73	16.81
G2	20.96	19.82	13.09	16.90	14.69
G3	37.38	41.36	38.45	42.69	40.22
P1	52.79	55.90	35.05	31.05	33.38
P2	6.36	6.76	3.76	3.00	3.44
P3	5.67	8.82	0.92	3.36	1.94
O	0.00	0.00	0.00	0.17	0.07
∑ avg.use	32.88	33.57	29.58	31.21	30.26

Average use of Netspeak elements quantity is 30.26. Figure 2 shows the distribution of Netspeak elements by academic years. It is expected the new generations of students to be keener to the use of Netspeak but there are many other factors affecting the amount of used elements, such as student's activity or the quality of the discussion. [10, 11] In generation 08/09 as many as 45.42% of students have actively participated in discussion, but the generation of 09/10 had a very high quality discussion, so there are slight discrepancies regarding the mean values used in analyzes of Netspeak in the last four generations of students. Netspeak elements appear approximately equally often in all generations. Netspeak elements at least appear in the generation of 2010/2011, mostly in the generation of 2009/2010. But the difference is negligible and amounts only to 3.99.

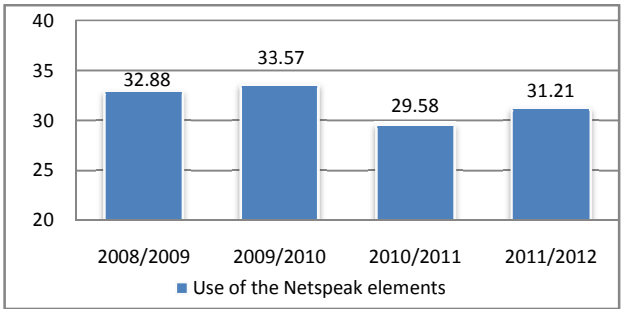

FIGURE 2
ARITHMETIC MEAN OF NETSPEAK ELEMENTS QUANTITY

To ensure that the arithmetic mean of the sample is representative and reliable indicator of the value of using Netspeak elements we have examined the standard deviation (Std. Dev) and the coefficient of variation (V) sample of students through the academic year. The results are shown in the Table 4.

TABLE 4
INDICATORS OF THE REPRESENTATIVE ARITHMETIC MEAN OF THE SAMPLE

	N	**Mean**	**Std. Dev.**	**V**
2008/2009	134	32.88	11.28	34.31%
2009/2010	110	33.57	14.19	42.27%
2010/2011	103	29.58	9.46	31.97%
2011/2012	74	31.21	11.16	35.76%

The results show that the coefficient of variation of the mean for all samples through academic year is less than 50%, which confirm that the arithmetic mean is representative enough.

Figure 3 shows the frequency of use of the each standard within the analyzed sample.

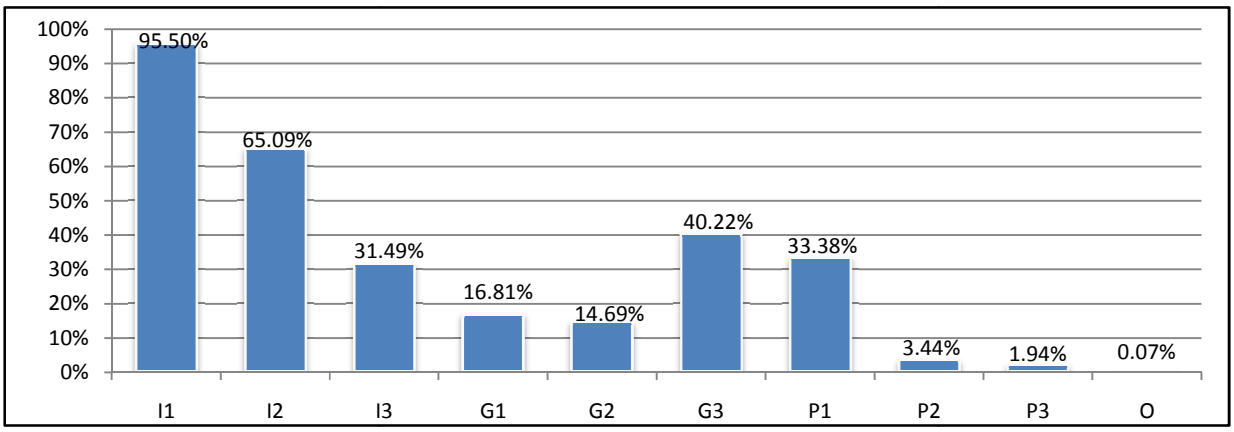

FIGURE 3
THE FREQUENCY OF USE OF THE EACH NETSPEAK STANDARD

As many as 95.50% of students use the standard I1 and that is words in English. Such a high percentage might be explained by the specificity of ICT course and the overwelhming use of English words for technical phenomena lacking the adeguate word in Croatian language. It can be assumed that within some other course this percentage might have been considerably lower. Also, 65.09% of students use standard linked to acronyms and abbreviations (I2). The standards G3, P1 and I3 in the second group according to frequency of their use. 40.22% of students ommit the diacritical marks (G3), while 33.38% of students use the punctuation in nonstandard way evoking some prosodyc effects for example, yelling, shouting, increasing or decreasing the intensity of the voice (P1), and 31.49% use emoticons (I3). 16.81% of students write the whole post using lower cases (G1), while 14.69% of students do not use space after the punctuation. Standards used below 5% are P2 concerning upper cases with 3.44%, standard P3 used 1.97% as well as the standard O considering the use of obsolete tense - aorist with just 0.07%.

Figure 4 shows the quantity of each standard within Netspeak elements used as a whole.

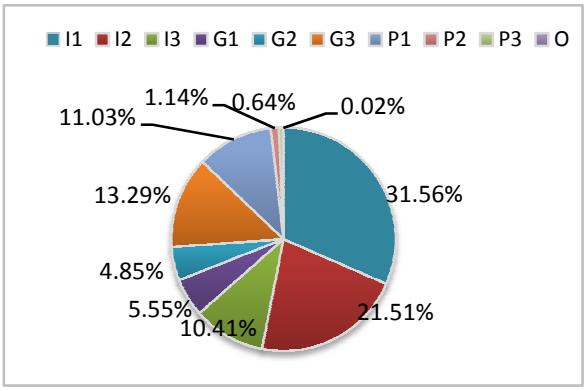

FIGURE 4
THE QUANTITY OF EACH STANDARDS OF NETSPEAK

31.56% refers to the standard I1, 21.51% refers to the standard I2, 13.29% to the standard G3, and 11.03% refers to the standard P1. Standard I3 is used 10.41%, G1 5.55%, G2 4.85%, P2 1.14%, P3 0.64%, while the standard O is used just 0.02%.

Figure 5 shows the distribution ratio by each group. Majority of the percentage, 63.47% goes to the group one, ICT; 23.7% belongs to the second group, Grammar; 12.81% is the percentage of standards belonging the third group, Prosody, and just 0.02% goes to others.

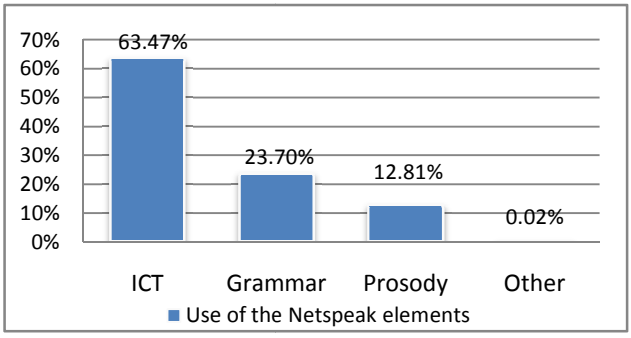

FIGURE 5
DISTRIBUTION OF STANDARDS WITHIN THE GROUPS

4. RESULTS ON STATISTICAL RESEARCH

The authors set the following hypothesis regarding the amount of Netspeak elements within the closed discussions:

1) There is a correlation between the uses of each Netspeak standards within the each group.
2) More active students or the students having the better quality discussions use more often the Netspeak elements then less active students.

4.1 Hypothesis 1 – Correlation between each standards within the group

Pearson correlation coefficient on the sample of 421 students shows how the use of elements are correlated as shown in the Table 5. Given statistics are the results obteined from „Paired simple t-test correlation“ analysis.

TABLE 5
CORRELATION WITHIN THE EACH GROUP

Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair 1	I1 & I2	421	.290	.000**
Pair 2	I1 & I3	421	.095	.053
Pair 3	I2 & I3	421	.149	.002**
Pair 4	G1 & G2	421	-.040	.413
Pair 5	G1 & G3	421	.117	.017*
Pair 6	G2 & G3	421	.100	.040*
Pair 7	P1 & P2	421	.164	.001**
Pair 8	P1 & P3	421	.187	.000**
Pair 9	P2 & P3	421	.123	.012*

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (1-tailed).

Within the ICT group there is a correlation between the use of English words and the use of acronyms and abbreviations and emoticons. Also, there is a correlation between the use of acronyms and abbreviations and emoticons. It is evident that there are statistically significant correlations between standards: I1 and I2, the correlation is significant at the 0.01 level and its value is 0.290 which is a weaker correlation. Standard I2, correlate with standard I1 as well as with the standard I3. In the group of Grammar there is a very weak correlation between using lower cases and the omission of diacritical marks as well as the low correlation between writing without using space and the omission of diacritics. Standard G1 correlate very weakly with the standard G3, with statistical significance of 95%. In weak correlation of 0.1 are also G2 and G3. Within the Prosody group there is a correlation between all standards, though it is very weak. The correlation ratio goes from 0.123 between the standards P2 and P3 to 0.187 between the standards P1 and P3.

4.2 Hypothesis 2 – Quantity of Netspeak and the quality of the discussions

In order to obtain as reliable sample as possible the authors analyzed the quality of discussions of the students who have participated in at least five different discussions. Quality of the discussion is measured with the method content analysis [10, 11]. The sample represents 104 students which is 24.7% of the whole sample that participated in the discussions. Table 6 proves that the chosen sample is representative.

TABLE 6
INDICATOR OF THE REPRESENTATIV SAMPLE

	N	**Mean**	**Std.dev.**	**V**
Use of the Netspeak elements	104	31.49	9.99	31.73%

Arithmetic mean of the selected sample of students coincides with the arithmetic mean calculated using Netspeak elements obtained by analyzing the sample of 421 students across all academic years, and the coefficient of variation confirmed that the arithmetic mean is representative.

Table 7 shows the correlation between the use of Netspeak elements and the quality of the discussions on the selected sample.

TABLE 7
CORRELATION BETWEEN MORE ACTIVE STUDENTS IN DISCUSSIONS AND THOSE WHICH THE DISCUSSIONS ARE OF BETTER QUALITY AND THE USE OF NETSPEAK

Correlations			
		The use of Netspeak	Quality of the discussion
The use of Netspeak	Pearson Corr.	1	.276**
	Sig. (2-tailed)		.005
	N	104	104
Quality of the discussion	Pearson Corr.	.276**	1
	Sig. (2-tailed)	.005	
	N	104	104

**. Correlation is significant at the 0.01 level (2-tailed).

It is possible to see that there is a correlation which value is 0.276, which with 2-tailed significance shows that there are weak links in the use of Netspeak elements and the quality of discussion. Or we can accept the hypothesis that students who are more active or have a better discussion more often used Netspeak than the less active students.

5. CONCLUSION

This paper develops ten standards in using Netspeak when communicating on-line and testes two hypotheses. The first one is that there is a correlation between the use of each Netspeak standard within the each group and the results confirmed this hypothesis. Although, only the use of words in English, use of acronyms and abbreviations and emoticons show high correlations while all others shows low correlation. This was expected due to the nature of the ICT course strongly oriented to overwhelming use of English language and its common feature of using acronyms and abbreviations. In terms of using emoticons in such a high rate it was also expected because symbols became very popular way of expressing a variety of feelings.

The results confirmed the second hypothesis as well and that is that the more active students or the students having the better quality discussions use more often the Netspeak elements then the less active students.

The standards are developed according to their provenience and grouped into four categories. The first one is ICT, considering standards that arise in the information communication technologies environment. The second one consists of standards linked to the grammar provenience and in the field of the third one arise the standards evoking prosody effects. The authors developed the fourth group, called "other" involving the use of an obsolete tense – aorist. The authors expect broadening of the standards as the new language will change and expand as well as the increase of the number of existing categories paralleling the growth of the new phenomena.

6. REFERENCES

[1] I.Tarnaj, "Impact and the function of media on globalization", University of Zagreb, Master thesis, 2004.

[2] A. Uzelac, "Impact of the new information communication technology on the cultural development: the role of the web", Doctoral thesis, 2003.

[3] K. Peovic Vukovic, "Text within the digital environment, third channel of Croatian radio", Assays and discussions, 2004.

[4] M. Mihaljevic, "Croatian language about Internet: Croatian language on Internet, **Language in social interaction**, (Stolac, D., Ivanetic, N., Pritchard, B.), 2005., pp. 319-330.

[5] K. Aleksic-Maslac, D. Vasic, J. Poropat Darrer, "Analysis of particular Netspeak elements in closed discussion within the Information Communication Technologies course", Proceedings: **Advanced educational technologies, 6th WSEAS/IASME International conference on Educational technologies (EDUTE '10)**, Kantaoui, Tunis, WSEAS PRESS, 2010., pp. 134-139.

[6] D. Vasic, K. Aleksic-Maslac, J. Poropat Darrer, "Impact of the Information Communication Technologies to the language changes and the creation of new language form – Netspeak", **EDEN 2010**, Valencia, Spain, 2010.

[7] K. Aleksic-Maslac, M. Korican, D. Njavro, "Important Role of Asynchronous Discussion in E-Learning System", **International Conference on Engineering Education and Research 2007** (ICEER 2007), Melbourne, Australia, 2007.

[8] K. Aleksic-Maslac, J. Poropat Darrer, T. Djuras, "Correlation and frequency of use of the Netspeak elements in asynchronous discussion within the same generation of students in the 1st and in the 7th semester", **WSEAS TRANSACTIONS on ADVANCES in ENGINEERING EDUCATION**, Issue 12, Volume 7, 2010.

[9] K. Aleksic-Maslac, M. Magzan, V. Juric, "Social phenomenon of community on online learning: digital interaction and collaborative learning experience", **WSEAS Transactions on Information Science and Applications,** Issue 8, Volume 6, 2009.

[10] M. Hammond, "A review of recent papers on online discussion in teaching and learning in higher education", **Journal of Asynchronous Learning Networks (JALN)**, Volume 9, Issue 3, 2005.

[11] A. K. Meyer, "Evaluating online discussions: four different frames of analysis", **JALN**, Volume 8, Issue 2, 2004.