

## Morphosyntactic Developments in Internet English: Statistical Analysis

Goda Rumšienė

**Abstract.** Internet English emerged and established itself during the final decade of the 20<sup>th</sup> century and it has been evolving into a universal means of communication on the world wide web. It has acquired a number of features generally common in pidgin languages, such as simplification of grammatical phenomena and regularisation. However, it manifests numerous specific elements such as a novel interpretation of word building processes, which emphasize a higher level of morphological complexity and thus tend to show differences from the patterns established in Standard English. On the other hand, simplification of syntactic structures within sentences is observed. Internet English is showing signs of logically-oriented developments as standardization of morphological structure, elimination of redundancies as well as meaningless elements together with simplification of syntactic patterns contribute to the facilitation of language mode acquisition by users who are not English native speakers. Multiple layers of contribution to the Internet language mode lead to extensive lexical innovations which nevertheless rarely deviate from the forms previously existing in the English language.

**Key words:** *Internet English, neologism, innovation, word/sentence length, word/sentence structure, pattern.*

### Introduction

Internet English and the Internet language in general evolved during the two final decades of the twentieth century. The difference in the way of communication in comparison with the so-called “standard” language appeared to be so large that the Internet type of communication was rapidly acknowledged to constitute the third medium of communication, which is different from spoken and written English (i.e. Standard English) in terms of grammar, vocabulary and type of interaction.

The culture of Internet English being largely affected by the specific environment of communication (interaction online and live, no prosodic or paralinguistic sign transmission possible, etc.), the online language developed towards extensive innovativeness in the areas of word formation, spelling and general expressiveness. The mode of language produced innovativeness as users stemming from multiple cultural environments aimed at gaining the highest amount of attention possible in the context of the English language providing the environment for intercultural communication, where the manifestation of emotionality represents a major part of creativity. The divergence from other modes of language is also manifested by referring to it as *netspeak* or *chatspeak* and thus by extension naming a member of the net community *a netizen*. These terms have also been widely used in the academic research on Internet language.

Generally, Internet English manifests numerous specific elements such as a novel interpretation of word building processes, which are frequently essentially different from the patterns established in Standard English. This is closely related to the fact that it is the third medium of interaction reflecting the general changes of relationships among cultures in the modern world, and the Internet is the cheapest and most readily available communicational environment whose use is restricted in the relatively small minority of censorship-practicing countries.

The process of the evolution of the Internet language started in highly civilized countries where Internet was available to larger amounts of the population. Not surprisingly, the technologically advanced countries, such as the United Kingdom, the USA and Australia constituted a major part of users of the Internet communicative environment. Statistics shows that with the universal spread of the availability of technologies, the decrease of the percentage of communication in English as well as sites in English is a natural phenomenon. Nevertheless, English remains the universal language of intercultural communication because of the general popularity of the language.

### Theoretical Background and Research Methodology

The exploration is based on the method of Technology-Conditioned Language Change and Use (TeLCU) by Adam Bodomo and Carmen Lee (Bodomo, Lee, 2002), on the methodology of Internet culture exploration by Constance Hale and Jessie Scanlon (Hale, Scanlon 1999) and on attitude to morphological analysis as expressed in the works by Martin Haspelmath (Haspelmath, 2003). Bodomo and Lee treat the phenomenon of Internet language as the modern manifestation of the communication culture, and, by taking the impact of technologies into account, claim that the new mode of communication also requires novel strategies of research. A similar attitude is taken by the author of the most famous monograph of Internet language, David Crystal (Crystal, 2001), i.e. that language evolves following the needs and achievements of its society. The paper takes *Ordinary Language Philosophy* by Mark A. K. Halliday (Halliday, 1985) as its philosophical concept; the ideas of this concept essentially overlap with the findings of Crystal and other leading explorers of Internet language. William Croft (Croft, 2005) among many others expresses the belief that the English language is rapidly evolving in terms of complying with the needs of the new type(s) of society, and this change is extremely prominent

in the medium of Internet English. In short, the predominant idea is that the lifestyle and technology of a civilization heavily impact upon its language and style of communication which in turn affects the civilization by impacting its lifestyle and development of communication-related technologies.

The *aim* of this research is to reveal the major trends of Internet English word and sentence structures in comparison with Standard English sentence structures.

The *tasks* of the research are as follows:

- To compare Standard English and Internet English in terms of word and sentence structure;
- To compare Standard English and Internet English in terms of lexical innovations;
- To demonstrate that the development of Internet English is not directed towards straightforward simplification of all domains of language.

The *object* of the research is Internet English and its word and sentence structure in particular. The materials employed in this research are the database of Internet English neologisms developed by the author of the present research and the database of neologisms of Standard English by Rice University, USA; the data from the year 2005 was employed, and the statistical generalizations cover the period of the years 1998 to 2005 inclusive. Internet neologism examples were taken from *dalnet*, *tsse* and *wow* chatrooms, and a number of general forums or forums covering various thematic areas. Lexical innovations first traced in the period of the years 2003 to 2009 inclusive are employed in this paper.

## Discussion

Internet English has been evolving into a universal means of communication on the World Wide Web. It has acquired a number of features generally common in pidgin languages, such as *simplification* of grammatical phenomena and *regularisation*:

bring→bringed, e.g. xnlm>damn it *bringed* me bad luck  
bad→baddest, e.g. snortel> thats the *baddest* we could get  
fast→fastly, e.g. browse photos and videos *fastly* on the internet, etc.

The above presented examples show cases of the shift from irregular or even suppletive forms to fully regularized ones, and, in the case of ‘fastly’, the standard adverb suffix *-ly* has been attached. On the other hand, no cases have been observed where traditionally regular verbs tend to acquire irregular forms.

Surprisingly, this phenomenon has occasionally been encountered among nouns as such forms as

box→boxen      fox→foxen

are used. However, this is rarely the case as these forms are used jokingly; they do not seem to oust the traditional variants *boxes*, *foxes*, etc.

The following types of differences in language use between Standard English and Internet English may be established with regards to the internationalization of the language:

- More complex word structure (increase in terms of morphemes per word);
- Decrease of the amount of functional words;
- Development and use of innovative vocabulary in the context of laxer application of the existing patterns and decreased number of bound morphemes;
- Decrease of sentence length;
- Systematization of language.

Having explored fragments of texts of Internet English (texts of conversations taking place from 2005 to 2007 in various chat rooms, such as *dalnet*, *tsse* and *wow*, were used) and Standard English (extracts of texts from “The Good German” by Robert Kanon and works by Margaret St. Clair, Cyrus MacMillan, Jack Denton Scott, Mari Sandoz, Jean McCord and Harold Coy) containing the same amounts of symbols (60,000 symbols each), the following results were produced (note that word form suffixes, such as *-s* or *-ed* were not included into the calculation, and e.g. *agreed* was treated as a one-syllable verb):

**Table 1.** Ratio of morphemes per word in Standard and Internet English

	Standard English		Internet English	
	Morphemes/ Words	Ratio	Morphemes/ Words	Ratio
Noun	2106/1542	1.366	2079/1477	1.408
Adjective	1144/738	1.550	1272/783	1.625
Verb	1420/1318	1.077	1516/1282	1.183

The above-presented table clearly shows that all three major types of parts of speech featuring affixational word building maintain the same trend: Internet English tends to use more derivational affix-possessing words. The comparisons of the ratios are also interesting:

1.408/1.366=1.0307 (for nouns),  
1.625/1.550=1.0484 (for adjectives),  
1.183/1.077=1.0984 (for verbs).

The Internet language exhibits a trend of the increased number of morphemes per word in nouns, adjectives and verbs. This data demonstrates relatively slight difference in terms of noun and adjective structure but major increase in bound morpheme(s) containing words. However, the attitude towards the morphological verb structure has undergone major alterations. There are two major factors to consider: first, Internet English actively employs verb formation suffixes *-en* and *-ize* (*to badden*, *to solariumize*, etc.), and second, Internet English introduces many verb compounds (e.g. *to copypaste*) while Standard English contains relatively few of them.

Internet English largely favours the substitution of complex phrases including articles and prepositions by complex words, for example:

Loads of cash→cashloads      On the spot→onspot

It is evident that the functional words “of” and “the” do not perform any function concerning contribution to the message; that is why they can easily be omitted by producing compounds. Internet English demonstrates trends of reducing phrases to single words, usually compounds.

If we compare the amount of functional words (by this we mean articles and some other determiners, e.g. any, prepositions, auxiliary verbs, etc.) as well as personal pronouns when the subject or object is evident from the context, the highest amount in the researched texts was 19% while in the researched Internet English conversations it never exceeded 12% and in occasional conversations was as low as roughly 5%.

In terms of sentence length calculation, *all* the words were considered. Standard English conversations were represented by transcripts of *Larry King Live* as posted on the official transcript page on the website of CNN. Internet English conversations were randomly selected from the above mentioned chat rooms.

**Table 2.** Average length of sentence in Standard English and Internet English

	Words/ sentences	Ratio (words per sentence)
Standard English	59,144/ 5,090	11.62
Internet English	56,827/ 6,871	8.27

The data clearly shows that Internet English sentences are substantially shorter, actually, the rate of difference is  $11.62/8.27=1.405$ , i.e. Standard English sentences are roughly 41% longer. Of course, this is not the absolute rate as registers in general as well as interlocutors in particular may substantially differ; however, Standard English sentences tend to have more ‘secondary’ parts and/ or more complicated structures than Internet English sentences. Generally, Internet English sentences do not present any systematic syntactic deviations; they tend to adhere to the same word order while differences mostly lie in the omission of inherently (contextually) clear subjects and objects.

In terms of ‘primary’ (subject plus object) versus ‘secondary’ (other parts of the sentence) elements in the sentence, Internet English exhibits a larger concentration of the ‘primary’ parts:

**Table 3.** Ratio of subjects plus objects per total words in a sentence in Standard English and Internet English

	Subject+object/total words	Percentage
Internet English	17,717/ 52,449	33.78%
Standard English	14,060/ 56,196	25.02%

The language clearly shows an inclination towards the decrease of sentence length. In chat conversations, users usually press the “enter” button in order to make a phrase appear on the screen. As conversations frequently hold more than 3 participants (this number may easily reach a few dozen), any sentence which is made longer than 8 words will be spatially too distant to the phrase it is intended to reply to. As a result, the message shows an inclination to be maximally short. Further researches could prove that 2-interlocutor conversations on the Internet tend to contain longer sentences than multi-participant interactions.

It is evident that major differences in terms of sentence structure have developed, the most prominent being the share of verb-less sentences, namely  $21.64/1.36=15.91$

times (i.e. 1,591 %), while verb-less sentences tend to be very short in Internet English. Further researches in this area should demonstrate that verb-less sentences in Internet English tend to be substantially shorter than 1 verb containing sentences.

**Table 4.** Distribution of sentences in Standard English and Internet English in terms of verbs per sentence

	Incomplete verb-less sentences	1-verb sentences	2-verb sentences	3+ verb sentences	Average verbs per sentence
SE	1.36%	67.59%	28.94%	2.11%	1.35
IE	21.64%	46.22%	29.98%	2.16%	1.12

However, the situation is more complex if we consider that Internet English features virtually the same percentage of 2- and 3+ verb sentences as Standard English. As a result, the share of 1-verb sentences in comparison with multi-verb sentences substantially increases with regards to this ratio in Internet English. This proves that Internet English conversations do not shy away from complex forms of expression. Actually, Internet English communicative style adopts rationalization and structural economy, but the ability of language to express several groups of ideas within a single sentence is fully preserved.

In conversations with no more than a few participants, phrases are frequently split, e.g.

xnbln>be back  
 xnbln >tomorrow  
 xnbln >6:30

This is visually acceptable, and it allows the initial parts of the message to get through faster.

It is evident that the style of word and sentence structuration in Internet English is substantially different from that of Standard English. Internet English develops more morphologically complex words which in some cases even represent parts of sentences and thus omits the structure-holding elements, such as articles and other functional words. Internet English sentences far more frequently feature the omission of the subject or the object, but the whole sentence structure remains the same as in Standard English, and the share of 2+ verb sentences is roughly the same in both variations of language. Due to the communicative environment of Internet English, the patterns of sentence layout may differ from the traditional single-line concept of Standard English.

In terms of neologism introduction during the period of eight years, from 1998 to 2005, the database of neologisms of Standard English by Rice University, USA attested 4367 neologisms, which is roughly 450 per year, 311 of which were difficult to attribute to a specific word formation type and thus were not considered in this research. It is inherently impossible to track the number of established Internet English neologisms, but the analysis of a relatively small amount of Internet communication, around six thousand pages of texts covering approximately two years, yielded 755 neologisms. Sources as [www.urbandictionary.com](http://www.urbandictionary.com) give reference to and provide definitions of tens of thousands of

lexical units, a large share of which are hardly encountered beyond the communicative interface of the Internet. As a result, it is very likely that Internet English introduces five or more times the amount of neologisms in comparison to Standard English. Unfortunately, such sources do not provide any specific filters which would make possible singling the Internet-restricted vocabulary out. Another issue is the unreliability of a large number of definitions as well as the presence of nonce-words. This largely devalues such user-updated neologism-featuring dictionaries, which thus may only provide a general impression rather than reliable scientific data.

In terms of derivational activity, first of all let us compare the relative frequency of various types of neologisms in terms of type of formation:

**Table 5.** Percentual frequency of types of word formation in Internet English

Type of word formation	Number of occurrences	Percentage of occurrences
Affixation	227	30.06
New root	68	9.01
Compound	221	29.27
Acronym	53	7.02
Clipping	41	5.43
Portmanteau	50	6.62
Word form alteration	43	5.70
Conversion	52	6.89
Total	755	100.00

Table 5 demonstrates the prevalence of affixation and compounding, which occur in more than a half of all the researched neologisms while the five remaining types are distributed more or less evenly and represent from 5 to 9 per cent of the share of total occurrences.

A very interesting phenomenon of word form alteration is to be singled out as part of the whole process of word formation. This does not lead to the shift of the meaning value of a word but is likely to be the first step of forming a new lexeme.

Generalizations grounded on the database by Rice University are represented in Table 6:

**Table 6.** Percentual frequency of types of word formation in Standard English

Type of word formation	Number of occurrences	Percentage of occurrences
Affixation	831	20.49
New root	992	24.46
Compound	510	12.57
Acronym	539	13.29
Clipping	703	17.33
Portmanteau	384	9.47
Conversion	97	2.39
Total	4056	100.00

Standard English clearly gives preference to new root and affixation neologisms, but the two leading types do not reach the mark of 50 per cent, and the remaining types show a broader distribution from 2 to 17 per cent.

The largest difference in terms of relative frequency is observed in new roots. It equals  $24.46/9.01=2.71$  times. If we consider that it is more or less the same people who use both of the modes of language but merely express themselves in different ways due to the peculiarities of the relevant communicative environments, the grounds of the communicational interface are the only source of the shift in the prevalence of some particular pattern(s).

It seems that the acceptance of new roots on the Internet is very low due to the reason that the users are separated spatially and thus also culturally. As a result, it becomes much more complicated to effectively spread a new root when few extra-linguistic signs are available to support the transfer of the meaning of the 'unknown' root. This is also affected by the presumption that Internet English users consider that they need to name only these objects which are reflected by corresponding lexical units in Standard English whose root stock is evidently adequate for the needs of Internet communication.

On the other hand, Internet English intensively develops shortenings, such as *iso* (instead of) or *imvho* (in my very humble opinion). This is related with their excessive popularity and consequently frequent occurrence as well as convenience of use and major economy of expression. However, the most important factor of their popularity is their belonging to the few phenomena that are believed to shape the style and identity of the medium of interaction. It is namely the correct use and complete understanding of the meaning of such shortenings that allow singling out the real partisans of the culture from the "outsiders".

Interesting data may be obtained by juxtaposing the frequency of multimorphemic lexical innovations. Thus if we compare the relative frequency rate of neologisms featuring the increase in the number of morphemes (namely, affixation and compounding), we get the following results:

**Table 7.** Percentual frequency of affix-based and compounding neologisms in Standard and Internet English

	Standard English	Internet English
Affixation	20.49%	30.06%
Compounding	12.57%	29.27%
Total	33.06%	59.33%

This data explicitly shows that internet English gives preference to morpheme-based word formation while Standard English exhibits a significantly higher relative frequency of new roots (roughly 2.6 times more) than Internet English.

While Standard English actively uses approximately 77 affixes for word building (cf. Crystal, 2003), Internet English is limited to roughly 30. This does not mean that Internet English is unable to express something but rather due to the fact that it does not tolerate redundancies and any rarer affixes performing overlapping or identical functions are ultimately replaced by the most frequent ones, i.e. *deferral*→*deferment*; there is no reason for the suffix "-al" to exist because there are more frequent affixes producing abstract nouns.

The only manifestation of the opposite process is the change of the status of some morphemes, i.e. some free

morphemes are so frequent in compounds that they may be treated as bound morphemes. Consider “ass” in:

Dumbass      Badass      Blackass, etc.

The recent developments in Internet English satisfy the major requirements which are set for languages used in multicultural and multilingual environments:

- Regularization of inflexion forms;
- Regularization of word building patterns by using relatively limited resources;
- Simplification of syntactic sentence-level structures;
- Simplification of some grammatical structures.

It is evident that such language developments render its acquisition easier and do not stifle creativity.

However, Internet English is not just “English made simple”. The major difference from a “merely simplified language” is:

- Many innovations frequently requiring deep understanding of language logic;
- Extensive active vocabulary featuring a large amount of roots;
- Preservation of the complex tense structure and frequent use of modal structures.

These developments facilitate the process of acquisition the language in general and mastering the skill of Internet interaction in particular as language tends to follow specific patterns, i.e. simpler sentences → easier to read and understand; standardized use of affixes plus usage of fewer morphemes → fewer “language brick” elements to learn plus easier to guess the meaning of an unknown word or to coin a new one.

## Conclusions

The following conclusions may be drawn:

- English is the dominant language of the World Wide Web. However, the share of Internet users employing English decreases in the context of the increase of the universal spread of the world wide web and the development of sites and other communicational

environments used within the limits of a single nation or culture.

- Internet English is a variety which is well-adapted to intercultural communication. It is simplified in comparison to Standard English in terms of grammar. In terms of lexis, very few new roots enter the Internet variety of English.
- Due to the contribution of those Internet English users who are not native speakers of English, the Internet language mode develops towards simplification and regularity. More and more patterns of word forms and word formation evolve. Internet English introduces roughly four times more novel vocabulary than Standard English.
- International contribution to the medium of the Internet language mode leads to multiple lexical innovations. However, lexical and flecional innovations rarely deviate from forms previously existing in the English language.

## References

1. Bodomo, A., Lee, C. K. M., 2002. Changing Forms of Language and Literacy: Technobabble and Mobile Phone Communication. *Literacy and Numeracy Studies* 12(1). Available at: <[www.readingmatrix.com/articles/bodomo\\_lam\\_lee/article.pdf](http://www.readingmatrix.com/articles/bodomo_lam_lee/article.pdf)>.
2. Croft, W., 2005. *Toward a New Theory of Language*. Available at: <<http://lings.ln.man.ac.uk/Info/staff/WAC/Papers/SanMarino.pdf>>. Not published. Posted as of 2005.
3. Crystal, D., 2001. *Language and the Internet*. Cambridge: Cambridge University Press.
4. Crystal, D., 2003. *A Dictionary of Linguistics and Poetics*. Oxford: Blackwell.
5. Hale, C., Scanlon, J., 1999. *Wired Style*. New York: Broadway Books.
6. Halliday, M. A. K. et al., 1985. *Language, Context and Text: a Social Semiotic Perspective*. Oxford: Oxford University Press.
7. Haspelmath, M., 2002. *Understanding Morphology*. London: Arnold.
8. *Neologism Statistical Database*. Rice University. Available at: <<http://esa4.rice.edu/~ling215/>>.
9. Rumšienė, G., 2007. *Word Formation Processes in Internet English*. Doctoral Dissertation. Vilnius University.
10. Shortis, T., 2001. *The Language of ICT*. London, New York: Routledge.
11. *Urban Dictionary*. Available at: <[www.urbandictionary.com](http://www.urbandictionary.com)>

Goda Rumšienė

## Morfosintaksinė internetinės anglų kalbos raida: statistinė analizė

Santrauka

Internetinės anglų kalbos žodžių ir sakinių struktūros esmingai skiriasi nuo standartinės anglų kalbos analogiškų struktūrų. Internetinei anglų kalbai būdingi morfologiškai sudėtingesni žodžiai, sudarantys santykinai trumpesnius sakinius, kurių didesnę procentinę dalį sudaro veiksniai ir tariniai. Vengiama funkcinių žodžių, kurie nėra būtini sakinio minčiai perteikti. Tarp naujų sutinkama beveik 80 % daugiau žodžių, kurių daryba paremta afiksacija ar šaknų sudūrimu. Tokios kalbos vartojimo nuostatos visų pirma paremtos ekonomijos ir racionalizacijos siekiu, artimai susijusiu su kalbos vartojimo aplinka. Daugiakultūrė internetinės kalbos vartojimo aplinka lemia išradimą vartojamų formų standartizavimą, o inovatyvumas yra paremtas originaliu jau egzistuojančių pavyzdžių interpretavimu. Analogais nesiremiam tik kuriant trumpinius ar emocijas perteikiančių simbolių kombinacijas. Šios stilistinės, sintaksinės ir morfolginės nuostatos patvirtina bendrąją nuostatą, kad internetinė kalba laikytina trečiąja kalbine erdve greta šnekamosios ir rašytinės kalbos, o jos raidos procese atsirandančios įvairių lygmenų struktūros nebūtinai įgis analogų kitose kalbinėse erdvėse.

Straipsnis įteiktas 2010 01  
Parengtas spaudai 2010 09

## About the author

**Goda Rumšienė**, PhD, Associate Professor at Kaunas Faculty of Humanities, Vilnius University, Lithuania.

*Research interests:* Internet English, sociolinguistics, verbal and non-verbal communication.

*Address:* Kaunas Faculty of Humanities, Vilnius University, Muitinės Str. 8, LT-44280 Kaunas, Lithuania.

*E-mail:* goda.rumsiene@yahoo.com

DOI: 10.5755/j01.sal.1.17.43380