Tanulmány

Ádám Porkoláb

"I'm tired of too professional, standard articles"

Gender-based analysis of young Hungarians' blogs

Abstract

"I like reading them [the blogs] because I always find personal writings there. I'm tired of too professional, standard news, articles… blogs are in a way different, and each is unique. They will never die out, at least I hope so." wrote Zoltán Bíró, a Hungarian blogger once about his own "profession". In this primarily qualitative paper, I try to point out how the differences between the communication styles of men and women appear in Hungarian blogs and I analyze the entries and the comments from a corpus linguistic perspective. I used a small sample, I analyzed the 5 latest entries and their comments on 20 blogs (10 written by male bloggers and 10 by female bloggers), altogether yielding 54,401 words.
Keywords: blogs, gender linguistics, net linguistics, Hungarian blogosphere, content analysis

1   Introduction

The literature on weblogs has provided an overview of the topic of blogs and blogging from a number of aspects since the early 2000's. However, these well-developed corpus linguistic aspects could not really take root in Hungarian internet linguistics. This is best evidenced by the fact that apart from a few isolated cases (Borgulya and Háhn 2013; Jóri 2012 and 2015; Petykó 2012, and from the perspective of advertisements Erőss 2009 and Háhn 2011) no corpus linguistic analysis has been produced about Hungarian blogs which would meet the requirements of modern linguistic discourse. The majority of works in Hungarian are concerned with providing a definition of the blog genre (Bódi 2010a-b; Csala 2005 and 2010; Juhász 2013; Petykó 2012).

According to the data of Statista.com[1], the number of blogs in the world shows a fivefold increase between October 2006 and October 2011. In 2006, the base year, we knew of only 35.8 million blogs, but by the end of 2011, there were 173 million blogs online. The popularity of blogs is even better illustrated by the fact that according to a different source[2] by July 2016 the cumulative number of Tumblr blogs was fourteen times higher than in May 2011. At that point in time 17.5 million Tumblr blogs existed, while 305.9 million such blogs were online in July 2016. I could not find data regarding the percentage of blogs written in Hungarian in the above cited source. Google returned 213 million hits for the term "magyar blog" ['Hungarian blog'] on August 11, 2016, but the actual number of Hungarian blogs –

---

[1]   http://www.statista.com/statistics/278527/number-of-blogs-worldwide/ (Accessed: 2016-09-15)
[2]   http://www.statista.com/statistics/256235/total-cumulative-number-of-tumblr-blogs/ (Accessed: 2016-09-15)

according to my own estimation – can only be a fraction of this, about 200-300 thousand. My estimation is based on the fact that 33,000 Hungarian blogs existed in August 2004, while in February 2005 this number was 82,000. (Bőgel, 2006: 34).

According to the data of the Wayback Machine website the first Hungarian language blog was a personal blog, titled "Technocol" which existed from 2000 until 2009.[3]

Concerning blogs, the international literature deals with their role in English teaching (Horváth, 2009), the analysis of comments on blogs (Mishne and Glance, 2006), their effect on the daily routines of bloggers (agenda-setting) (Delwiche, 2005), the social network of bloggers and users writing comments (Marlow, 2006), the propagation of topics and information on blogs (Gruhl, Guha, Liben-Nowell, & Tomkins, 2004) and blogging as a social act (Miller and Shepherd, 2004). The importance of researching blogs was recognized as early as 2003: Kumar and Novak (2003: 1) point out two important reasons to study the blogosphere. The sociological reasons focus on the more restricted, local aspects of the blogosphere, whereas the technical reasons concentrate on the relatedness and temporal change of traditional internet research and the blogosphere.

The main aim of this paper is to analyze the budding Hungarian blogosphere through qualitative methods, primarily focusing on the sociolinguistic category of gender, and to compare the results with the findings published in the topic's international literature**.**

2 Definitions of blog and features of the genre

2.1 A brief history of blogs

According to the literature, the first web diary which could be considered a blog was created by Tim Berners-Lee in 1991 (Winer, 2002). The prototype of modern blogs was created in 1996 by Dave Winer as part of 24 Hours of Democracy[4] (Herring, Scheidt, Bonus, & Wright, 2004). The term itself was first used in 1997 by Jorn Barger (originally in the form *weblog*), the term *blog* first appeared in 1999 (Herring, Scheidt, Bonus & Wright, 2004: 1).

The advent of blogs are the result of a number of factors: from the perspective of media history, 9/11 and the Iraq War had positive effects on blogging, because they provided a new platform for dialogue and made communication faster (Khrishnamurthy, 2002). Besides, in 1999 a specialized software was released which helped the publication of blog posts (Gruhl, Guha, Liben-Nowell, & Tomkins, 2004, 49). Since the terrorist attacks on September 11, 2001, blogs have been a tool of personal journalism. Delwiche, to substantiate this claim, quotes a number of authors who call blogging "pamphleteer journalism", "amateur journalism" or "do-it-yourself journalism" (Delwiche, 2005: 3).

From the perspective of genre history, McNeill (2005) found a correlation between blogs and traditional diaries, thereby disproving the claim that the genre of blog is one without any precursors. According to Miller and Shepherd (2004) the blog's evolution was conditioned by present-day needs ("the blog is an evolutionary product" – Miller and Shepherd, 2004: 3). These authors also discuss that the blog genre is determined by the temporal features of the present day (they use the Greek term *kairos*, which means appropriate, adequate, suitable moment).

---

[3]   Due to the restrictions of length, I cannot present a history of Hungarian blogs here. For a thorough summary see Erőss (2009: 17-23).

[4]   On 15.09.2016 the webpage was still available at http://scripting.com/twentyFour/default.html

2.2   Definitions of the blog

Given that the early literature dealing with blogs (1999-2005) focused rather heavily on providing a genre definition, today we can define the phenomenon with appropriate adequacy. Susan Herring and her colleagues, widely considered to be experts in the field, give the following definition: "frequently modified web pages in which dated entries are listed in reverse chronological order" (2004: 1). Zoltán Bódi – one of the first Hungarian researchers of internet linguistics – (2010b) also provides a definition based on Herring: "The weblog is a website or webpage which contains short, frequently updated personal entries intended for the public to read. The entries are posted in a chronological order" (2010b: 37). According to Juhász: "blogs are frequently updated webpages where the entries are connected to dates and are editable going forwards or backwards in time" (2013: 4).

Petykó, at the end of his corpus-based cognitive linguistic analysis defines the genre of the online diary as follows:

> a blog can be defined as a dynamically changing series of discourses published on the internet under a public address, which are dated, follow each other linearly and contain at least one discourse-initiating entry (post) which functions as a first and unique utterance (2012: 65).

David Crystal defines the nature of blogs in the following way: "It takes the form of a personalized web page where the owner can post messages at intervals" (2006: 15, cf. Crystal, 2011). Crystal emphasizes the unedited nature of blog entries (unmediated). It is also him who calls attention to the new words which emerged in the blogosphere: blog hopping, bleg ('begging blog'), blogorrhea, bloggerina ('female blogger'), blog swarm, etc (2006: 239).

Miller and Shepherd's definition (2004: 15) suggests a different approach: "We see the blog, then, as a genre that addresses a timeless rhetorical exigence in ways that are specific to its time".

I define a blog as any regularly updated webpage consisting of entries usually in reverse chronological order, which is written as a blog as per the declared intention of the author or authors. In my opinion, the main purpose of the blog is no longer to share personal experience and information but the social utility and marketing possibilities should also be taken into account.

2.3   The types of blogs

Despite the apparent uniformity of blogs, which result from their popularity, the genre is suitable for achieving a surprisingly wide variety of communicative goals. One of the first tasks of research in the literature was – besides gathering the attributes characteristic of blogs – to create a typology of the genre.

Blood was the first to describe (2002) the archetypes of blog genres. She originally only distinguished between two types: the filter-style and the more personal blog-style.

Krishnamurthy (2002) categorizes blogs along two dimensions: personal and topical. According to these criteria, Krishnamurthy (2002) differentiates online diaries, support groups, enhanced columns and collaborative content.

In their study, Miller and Shepherd (2004: 7) outline the thematic categorization of The Weblog Review site: adult, animation, camgirl, computer, entertainment, humor, film, music, news/links, personal, photo, Spanish/Portuguese, teen and video games.

Herring, Scheidt, Bonus & Wright (2004) categorize blogs according to their function. They build their research on Rebecca Blood's (2002) categories.

- Personal blog (personal journals)
- "filter-style weblog" (describing events external to the author's life)
- Knowledge-log, k-log for short (collecting information related to a specific topic)

One of the first researchers of the topic in the Hungarian literature was Bertalan Csala (2005 originally, then revised in 2010) who groups blogs according to formal criteria as follows:

- Text-only blogs (very rare)
- Photoblogs
- Link-collecting blog
- Personal blog
- E-media (blogazine)
- Community blog
- Thematic blog
- Blog comics

Judit Bayer (2005: 15) provides the following definition: "Diary-like, frequently posted, chronologically ordered entries, written by one or more persons."

My personal opinion is that the majority of such categorizations focus solely on the form and content of blogs and – similarly to Miller and Shepherd's (2004: 7) definition – do not see blogging as a social action. I propose the addition of a new criterion besides the traditional categories, which can successfully factor in the effects of new trends and fads. The way I see it, a different, more universal type of categorization is needed, one which is built primarily on social functions and is less sensitive to the developments of blogging software (i.e. one that is not based on formal criteria and authorship). In this framework I distinguish two major groups of blogs:

The first group is connected to the classic function of the blog[5]: in this thematic group the blog is the tool of self-publishing. Here blogging is done in favor of an individual or a smaller group, they are the main benefactors (in the cultural, recreational and financial sense of the word).

I categorize personal blogs into this group, as well as any other blog-like content which aims at the publication of literary or scientific works and personal thoughts or provides a platform for the voice of political opposition. Blogs promoting the work of various organizations, foundations and companies also belong to this category.

In the other group, blogging serves the benefit of society or the need of a larger group. Blog magazines, consumer protection blogs, and online diaries written for sociological, pedagogical or anthropological purposes belong here. This group comprises those contents which are not primarily produced for self-publishing.[6]

Undoubtedly, this categorization takes profit-orientedness into account, but with this I would like to point out another phenomenon: by the 2010's it has become common practice for larger companies in the beauty industry and in the IT industry to build on the advertising potential of popular blogs. In the case of Hungarian blogs – especially for the more popular

---

[5] Cf. Miller and Shepherd, 2004, 6: "*It appears that blogs originated as a way to share information of interest*".

[6] Naturally, I do not claim that any blog can be neatly and unambiguously fitted into one category or the other. These two groups will definitely need to be refined and expanded later, because, for example, editors of blog magazines are remunerated for their work which seemingly contradicts the definition of the group.

women's blogs – product tests requested and sponsored by beauty companies have become widespread, as well as marketing-oriented events organized for bloggers and vloggers.[7] Recently, beauty and fashion blogging have become the focus of much internet linguistic research (for instance see Marwick 2013, McQuarrie, Miller & Phillips 2013, Mora & Rocamora 2015, Ots & Abidin 2015, Pihl 2014, Tan 2016), and I anticipate that these topics will be picked up by Hungarian scholars too in the near future.

Blogs and their authors have, without doubt, become brands and new advertising platforms for Generations Y and Z: all this inevitably foreshadows the decrease of personal content in the future which will eventually give way to sponsored product tests and news from various festivals and events. The emergence of Web 2.0 completely transformed the traditional format of blogs. These pages are becoming more and more hybridized: from the analysis presented in later subsections of this paper it clearly emerges that online diaries are tightly interwoven with content sharing platforms and the social media. Blogs are no longer used for publishing chiefly textual content, but have become link collections for various vlog videos.

The fact that marketing infiltrated blogs completely reevaluates the feature of blogs that Peter Elbow (quoted in Miller and Shepherd, 2004: 9) calls *private writing*: the aim is – with the monetization of Youtube and other content providers – no longer just the production of personal content without the fear of the reaction of others. The role of blogs is different today: it is to shape the world view and needs of the youth. Comments and likes provide feedback for users and therefore are highly important in this process. Interactions taking place within blogs have different names in the literature: "conversations", "communities", "bursty interactions" and "political debate" (based on Marlow, 2006: 1).

3   Research questions and the corpus

While planning the research, it became apparent that carrying out a large-sample, quantitative analysis, as favored by the literature, would not be possible with the materials that I had at my disposal. Moreover, such an approach would not be appropriate for the purposes of the current investigation. The weakness of this paper is that at the current stage of research it only allows the analysis of binary gender constructs. Future research will have to be extended to include the analysis of blogs by members of the LGBTQ community (lesbian, gay, bisexual, transgender, queer), but – in lack of reliable Hungarian data – such an approach cannot be followed in the present paper. The classification of the bloggers' gender is based on self-assessment (e.g. civil name, narrative of blog entries).

The research questions for this study below were formulated after surveying the literature on the general, linguistic properties of blogs and on the speech styles of the genders.

1. To what extent can the statements of the English language literature about gender be applied to the Hungarian blogosphere? My assumption is that there will be a significant difference between male and female bloggers and I believe it is likely that the mode of keeping in touch with the readers is not going to be the same between the two genders.

2. From a corpus linguistic perspective what differences can be observed in the two subcorpora organized by the blogger's gender? Which parameters can be used for determining the blogger's gender?

---

[7]   Since 2012 the most popular vloggers in Hungary can meet with their fans at the Online Video Stars festival.

3. Will the commenting users' role be the same in both subcorpora and will they follow the online etiquette to the same extent? Given that the authors of comments cannot be filtered by their gender, I presumed that this analysis will not reveal significant differences.

My aim was to scrutinize the 5 latest entries and their comments on 20 blogs (10 by male and 10 by female authors) primarily selected on the basis of the gender of authors (altogether 100 posts) and to perform a deep qualitative analysis on the sample.[8] The entire corpus consists of 54,401 words (tokens) and 15,184 word-forms (types). The type-token ratio of the corpus is 0.279. Concerning their distribution, the blog posts on average cover 34 days.

Methods of classical sociolinguistics, gender linguistics, and corpus linguistics were utilized for the analysis, involving data from authors of different genders from a number of perspectives, which include the "age" of the blog itself, the frequency of posts, the average length of posts, the number of comments, the appearance of jargon in comments and in posts, the number of internal links (self-quotes) and external links. In addition, I also took into consideration the number of anonymous comments. I laid down these aspects while bearing in mind the issues already explored in the literature, chiefly because to my knowledge no analysis has yet been done on the Hungarian blogosphere which would be as thorough as the present one. Hopefully my work can, to some degree, contribute to the research of the non-English language blogosphere and can inspire further research in the area. My conviction is that – similarly to gender linguistic research – the research of the blogosphere requires an international foundation.

3.1   Criteria of inclusion and contents of the corpus

The 100 posts gathered from the blogs can be divided into two subcorpora: the first half of the corpus – that is, 50 entries – is comprised of the first five posts of ten beauty blogs. The authors of the entries are all women. The blogs were collected from the Magyarblog.hu aggregator site's "beauty" category based on the database of 06.08.2016.

The 50 entries of the male subcorpus were collected from blogs in the "Tech" category of the Blog.hu aggregator site (http://blog.hu/rovatok/tech), based similarly on the database of 06.08.2016.

The corpus of Hungarian language posts was compiled according to the following principles: blogs that only had a handful of readers or had hundreds of thousands of page views were not included. The selected blogs had to represent as wide a range of hobbies and activities as possible. The blogs selected this way cover the following topics and hobbies: beauty treatment, reading, do-it-yourself (DIY) activities, science and technology, sound systems, board games, video games, action figures, and cookery. This mode of selection was not employed to serve quantitative research methods, which means that the aim was not to collect as large and diverse a sample as possible.

I removed all the recurring lines that are not produced by users (e.g. "Reply", "Terms and Conditions", permalink, etc.) and the filtered, non-lemmatized corpus was edited and stored in a word processor. I used the KWIC Concordance for Windows software, version 5.1.0[9] for the

---

[8]   For comparisons: Mishne and Glance (2006) analyzed 645,000 comments, Delwiche (2005) analyzed 781 news stories, Petykó (2012) 500 discourses of 100 blogs, and Marlow (2006) 5000 webpages.

[9]   The freeware program was created by Satoru Tsukamoto. With its help one can perform concordance analysis, compile a word list, and search for keywords within the corpus which are displayed within context. Currently (in October 2016) the latest version is 5.3. The software can be freely downloaded without restrictions from http://www.chs.nihon-u.ac.jp/eng_dpt/tukamoto/kwic_e.html.

corpus linguistic analysis. The main features of the male and female subcorpora are shown in Tables 1 and 2.

Name of blog	*Combined length of entries (characters, including spaces)*	*Average length of entries (characters, including spaces)*	*Number of comments*	*Number of multimedia content (picture and video)*	*Tags (number)*
Azrail	8669	1734	37	21	18
Beautycrumbs	6099	1220	23	19	15
Breakfast at Fruh	7285	1457	4	11	47
Csipkelány ('Lacegirl')	12227	2445	19	35	9
DJ Katy Beauty	14765	2953	44	25	64
Eyeliner and smile Andi	14506	2901	10	30	16
Fanni Gardróbja ('Fanni's Wardrobe')	13251	2650	50	5	16
NJoyStyle Reni	16879	3376	16	74	0
Stílus és egyéb katasztrófák ('Style and other catastrophes')	5052	1010	45	11	12
The very style	12012	2402	24	11	15
Average:	11074	2215	27	24	21

Table 1: Main features of the female subcorpus

Name of blog	Combined length of entries (characters, including spaces)	Average length of entries (characters, including spaces)	Number of comments	Number of multimedia content (picture and video)	Tags (number)
Appleblog	8776	1755	32	6	6
Audioworld	52528	10506	0	39	9
AVGN és NC Rajongói oldal ('AVGN and NC Fan page')	900	180	17	4	4
Freedee	5029	1006	0	7	15
IDDQD	16627	3325	67	10	15
Munkakerülő ('Lazybones')	1712	342	2	5	12
Pontehu	5095	1019	0	0	115
Star Wars Miniatures	23880	4776	0	42	24
Telefonközpont ('Phonecenter')	14547	2909	6	7	27
Vírusirtó blog ('Antivirus blog')	14862	2972	4	10	11
Average:	14396	2879	13	13	24

Table 2: Main features of the male subcorpus

I used the word counter of the word processor to determine the average length of entries, excluding comments from this analysis, concentrating on the raw text of the posts.

The male subcorpus contains 128 comments, 130 pieces of multimedia content and 238 tags, while the female subcorpus contains 272 comments, 242 pieces of multimedia content and 212 tags. In the case of the female blogs in the corpus, the number of authors per blog is usually 1, while in the male subcorpus 2 or 3 authors frequently collaborate in writing a single blog. Concerning the entirety of the corpus, the average age of the blogs is 1374 days, that is 3.76 years.

3.2   Corpus linguistic analysis of posts

The first thing that was considered while surveying the blog entries was the quantity of posts, the numbers indicated in the above tables have been rounded when needed. The corpus of male bloggers was richer in this respect: the average text length per blog in the subcorpus was about 3000 characters more than in the female corpus. However, this is a misleading figure, because through closer investigation it emerged that the length of posts in the male subcorpus shows an uneven distribution: the shortest and the longest entry are both found in the male corpus.

The male bloggers included in the investigation are more prone to rhapsodic alternations of post length. The shortest item in the corpus comes from the "AVGN és NC Rajongói oldal" ('Angry Video Game Nerd and Nostalgia Critic Fan page') blog: it consist of only 99 characters including spaces. The longest entry, consisting of 15,082 characters including spaces, comes from the "Audioworld" blog.

Based on the data I collected, a clear dominance of textual content within posts is observable in the case of the male subcorpus, however the images inserted into the entries are mostly made by the poster. Third-party images and embedded vlog videos occur more frequently in the female subcorpus. The reason behind this is that a number of female bloggers in the corpus are also active video bloggers and they share the links to their videos on their blog.

While in 2006 Crystal (referencing Wilbur)[10] claimed that spoken language was only marginally present on the internet, in 2016 the dominance of text-heavy material (including blogs) on the internet is not so obvious. During my analysis it emerged that one third of the entries in the female corpus contain videos and in other posts the amount of images show a clear dominance over the amount of text. Even if this finding fails to disprove Wilbur's claim, we still have to bear in mind that in the near future the dominance of textual material on blogs will decrease to some extent. However, this could only be confirmed or disproved by a fine-grained longitudinal analysis.

The issue of the authorship of the entries is not yet settled convincingly. It is debatable whether to consider links to music sharing websites (YouTube, Soundcloud) as own content if the link points to third-party content. Along with legal issues, this also raises the question of the genuine authorship of posts. If we consider sharing links with the readers to be a task related to the function of blogs then – from a linguistic and narrative aspect – these can be by all means appropriate. However, such investigations can be problematic and perhaps no such methodology exists within linguistics that could be applied here.

3.3   The frequency of posts

During the analysis of the entries I also examined their frequency. It emerges that in the female corpus on average a new post is published every 4.26 days, while in the male subcorpus the frequency is one new entry per 9.56 days. When these values are compared with the length of posts we can say that male bloggers post entries which are published five days later than entries in the female subcorpus and are 3000-3300 characters longer than items in the other corpus. In their corpus linguistic analyses, Gruhl, Guha, Liben-Nowell, &

---

[10] *"Whatever else Internet culture may be, it is still largely a textbased affair"* (Wilbur, 1996, 6 quoted in Crystal, 2006, 9).

Tomkins (2004: 493) found that the lowest number of posts are published on Sunday and – concerning the distribution of posts – Wednesday is the most productive day.

The picture can be further refined if we examine the temporal distribution of the 5 entries in each subcorpus. The 5 posts in the female subcorpus were published over an average of 21.3 calendar days, while the entries in the male corpus were put online with less frequency, over 47.8 days. Based on these pieces of data it seems logical to conclude that in the case of this corpus the number of posts is inversely proportional to the length of the posts, however I do not consider this phenomenon generalizable based on a small sample.

3.4   Internal and external links in the posts

It is a commonplace in gender studies that women have better social skills and they perform extra work to keep the conversation going.[11] Such analyses in gender studies focused primarily on verbal communication, but these phenomena can also be observed in online communication. Susan Herring's (1994) research was perhaps the first investigation of the difference of forum posts written by users of different genders. According to Herring, female communication is characterized by cooperative, polite, non-abusive (no flaming) and more reserved linguistic behavior (Herring, 1994). As opposed to this, men post more frequently, their tone is rather sarcastic, sometimes strong, and even contentious (ibid).

Therefore, it stands to reason that we will find different strategies in blogging utilized by representatives of different genders. One possible method of analysis is to survey the links on blogs from the perspective of their function. Depending on their direction, these can be categorized as *internal* or *external* links. I considered links pointing to other posts within the same blog self-promotion and self-marketing because they take the reader to another post by the same author. No significant difference emerged between the average amount of internal links in the two subcorpora: in the female sample the average number was 2.4 while in the male sample it was 3.

In the case of external links there is a duality in the analysis: on the one hand we can examine the external links to sites of partners, friends and sponsors placed on the blog's main page. I considered placing links like these to be strong social actions (except in cases when the links pointed to the author's social media page), because this way the blogger can popularize their peers which can even result in financial gain for the other party. In the case of women, the average number of external links was 2.7, while in the case of men it was – surprisingly – 3.9. This means that the assumption of women having better social skills cannot be verified based on the analysis of blog entries and blog home pages alone.

Besides the main page, external links can also be found within blog posts. Concerning the whole of the corpus, the blog entries contain on average 8.4 external links. The average amount in the female subcorpus is 6.9 while in the male subcorpus it is 9.9. The 10 blogs written by women altogether contained 69 external links, and the blogs written by men contained 99. The results are summarized in Table 3 below.

This method of analysis therefore can neither confirm nor disprove the classic thesis of gender linguistics about the social skills of women, moreover it appears that men in the sample produce longer entries and use more internal and external links. It seems then – based

---

[11]   See also: Huszár's (1994) co-operational coefficient, and Fishman's (1980) notion of "interactional shitwork".

on the links on social media sites and the amount of comments – that the amount of links is chiefly correlated with the length of the published text and not with the blogger's gender.

Name of blog	Average frequency of entries (days)	Distribution of entries (in calendar days)	Current number of authors	Age of blog (days)	Number of external links on blog	Number of external links in posts	Number of internal links
Appleblog	0.4	2	2	2754	4	9	0
Audioworld	5.4	27	2	185	0	1	0
AVGN és NC Rajongói oldal ('AVGN and NC Fan page')	41.4	207	1	1980	1	1	0
Azrail	1.4	7	1	1002	1	12	1
Beautycrumbs	2.6	13	1	1331	2	2	1
Breakfast at Fruh	8.4	42	1	1687	3	2	0
Csipkelány ('Lacegirl')	2.4	12	1	1322	1	3	2
DJ Katy Beauty	9	45	1	1288	0	41	15
Eyeliner and smile Andi	3.2	16	1	656	0	3	0
Fanni Gardróbja ('Fanni's Wardrobe')	10.2	51	1	1677	10	0	0
Freedee	0.6	3	2	926	9	7	1
IDDQD	1.8	9	3	2133	7	30	22
Munkakerülő ('Lazybones')	1.8	9	2	1888	4	5	0
NJoyStyle Reni	1.6	8	1	1090	7	3	2
Pontehu	1	5	1	1846	3	3	0

Star Wars Miniatures	3	15	2	179	6	35	7
Stílus és egyéb katasztrófák ('Style and other catastrophes')	1.8	9	1	1678	1	1	0
Telefonközpont ('Phonecenter')	36.4	182	1	1321	3	1	0
The very style	2	10	1	1255	2	2	3
Vírusirtó blog ('Antivirus blog')	3.8	19	1	1283	2	7	0
Average	6.91	34.55	-	1374	3	8	3

Table 3: Details of the blog entries and links

I should continue the examination of external links. This time I looked at the main pages of the blogs, aiming to collect data on the number of social media profiles associated with the blogger. The results are summarized in Table 4 below.

Name of blog	*Facebook*	*Twitter*	*Youtube*	*Instagram*	*Blog-lovin*	*Pinterest*	*Other* [12]
Appleblog	+	-	-	-	-	-	+ Em
Audioworld	+	-	-	-	-	-	-
AVGN és NC Rajongói oldal ('AVGN and NC Fan page')	-	-	-	-	-	-	+ Em
Azrail	+	-	-	+	+	-	+ Em, G
Beautycrumbs	+	-	+	+	+	-	+ Em, G, G2, W

---

[12] Abbreviations: Google Contact (G), Google Plus (G2), Newsletter (N), e-mail (Em), "We heart it"-profile (W)

Breakfast at Fruh	+	+	+	+	-	+	+ Em, G
Csipkelány ('Lacegirl')	+	-	+	+	-	-	+ G, Em, N
DJ Katy Beauty	+	+	-	+	+	-	+ G, N, Em
Eyeliner and smile Andi	-	-	-	-	+	-	+ G
Fanni Gardróbja ('Fanni's Wardrobe')	+	-	-	+	+	-	+ Em
Freedee	+	-	-	-	-	-	-
IDDQD	+	-	-	-	-	-	+ Em
Munkakerülő ('Lazybones')	-	-	-	-	-	-	-
NJoyStyle Reni	+	+	+	+	+	+	+ Em, G, G2
Pontehu	+	-	-	-	-	-	+ Em
Star Wars Miniatures	+	-	-	-	-	-	-
Stílus és egyéb katasztrófák ('Style and other catastrophes')	+	-	-	-	-	-	+ Em, G
Telefonközpont ('Phonecenter')	+	-	-	-	-	-	-
The very style	+	+	+	+	+	+	+ G

Vírusírtó blog ('Antivirus blog')	+	-	-	-	-	-	+ Em
Total:	17	4	5	8	7	3	17

Table 4: Number of social media profiles linked on the main page of blogs

The results indicate that the authors in the female subcorpus provide links to 4.9 different social media profiles on average on the main page of their blog. The most popular of such social media outlets among female bloggers is Facebook (9 bloggers use it) and Instagram (8 users) and Bloglovin' (7 users). In the case of the male subcorpus the average number of contact options given is 1.3.

From this aspect, the idea that women have better social skills in the domain of blogging could be proved, because they provide more platforms where their readers can contact them and leave feedback. However, we must also take into consideration the results of other researchers. Kumar and Novak (2003: 5) interpret the interconnected network of blogs in the completely different light of the state of affairs of 1999: "there is no strongly connected component of more than a few nodes". Marlow also calls attention to the fact that online relations break up more easily than personal ones. Acquaintances formed on the internet can only be strengthened through personal contact (Marlow, 2006: 2).

3.5 Other linguistic properties of the blog entries

Besides the frequency of publication, blog entries can also be analyzed from various linguistic perspectives. Concerning the language and the style of the posts, no significant differences are observable between the two subcorpora. Both use the same mixture of journalistic and personal style, and each entry contains on average 6-8 elements of jargon[13]. Obscene words only occur sporadically. All bloggers observe the norm of written language, spelling mistakes rarely occur.

However, one important difference has to be mentioned concerning the goal of communication: in the case of the female subcorpus, personal narratives and sharing experience is dominant. In blogs written by men, on the other hand, sharing facts and commenting on official news stories are of primary importance, which might be the reason for the higher number of external links in such entries. The nature of these links is also different, in the female subcorpus they mostly point to blogs of the author's friends, while in the male subcorpus the links point almost exclusively to official reports and statistics and websites of companies, thereby in a way supporting the blogger's credibility.

An important indicator of the content of entries is the conclusion at the end of the posts. Seven blogs in the female corpus contain questions at the end of posts. This indicates that a given blog has a well-established readership, because the author is aware of the fact that her readers know she addresses them and that they will leave comments. This is a social action –

---

[13] Examples from the male subcorpus: quadrocopter, space drone, platform level, X-Wing community, 5G telecommunication networks.
Examples from the female subcorpus: hydration, mechanical stress, ISO-value, beauty blender, APTYL 100 molecule complex.

relating to the discussion in Section 2.4 above – which is also associated with women, confirming the classic thesis of gender linguistics. Moreover, I must add that in certain cases of addressing the readers ("capitatio benevolentiae") cannot be analyzed from this aspect because those serve the interest of the blogger. (for instance: "Follow me on Twitter!", "Vote for me in this contest!") This is also a specific feature of the female corpus only.

The male corpus is consistent in this respect: apart from a few blogs[14] ("IDDQD Blog", "AVGN és NC Rajongó oldal" ['AVGN and NC Fan page']) they usually do not acknowledge the presence of the readers, they are not addressed directly. The reason behind this could be the fact that the corpus of blogs written by men contains half as many comments as the corpus of female bloggers.

3.6   Analyzing the comments

The corpus of texts can be analyzed from yet another perspective, which would be the comments. Altogether the corpus contains 400 comments, of which 128 come from the male subcorpus and 272 from the female one. Comments can be divided into two types, based on the role they play in relation to the discussion: ideal and not ideal comments. According to my definition (Porkoláb 2015) comments are ideally short messages reflecting the user's own personal opinion which is related to the content of the original post and is expressed in accordance with the appropriate rules of the environment. In online socio-culture most comments which violate the system of norms ("netiquette") are referred to with a special name by users. A typical violation of norms is when a user posts a self-reflecting first comment or when a user posts a series of comments unrelated to the original thread ("off-topic", "flooding") or when a user deliberately offends another user ("trolling").

Anonymous users are quite interesting from the point of view of online etiquette. In 1995, Section 4.1.2 of the Netiquette Guidelines (RFC 1855)[15] clearly accepted anonymous chatting, but since then the users' attitudes towards anonymity has changed. Since the appearance of Facebook, users have to make comments under their real names and anonymous users are almost automatically stigmatized as trolls and deviants (cf. Bayer, 2005). In my analysis I could not confirm whether or not this is true: in my corpus there are 22 anonymous comments but none of their authors violated any norms. Results are shown in the Table 5. below.

Name of blog	Total number of comments	Ideal comments from readers	Number of replies by the blogger in the comments	Anonymous comments
Appleblog	32	**25**	3	0
Audioworld	0	-	-	-

---

[14]   However, these instances of feedback are not so prominent as in the other group, they are not emphasized typographically by formatting or line spacing.

[15]   *"If a user is using a nickname alias or pseudonym, respect that user's desire for anonymity. Even if you and that person are close friends, it is more courteous to use his nickname. Do not use that person's real name online without permission."*

AVGN és NC Rajongói oldal ('AVGN and NC Fan page')	17	16	1	0
Azrail	37	28	9	2
Beautycrumbs	23	**22**	0	0
Breakfast at Fruh	4	3	1	1
Csipkelány ('Lacegirl')	19	19	0	3
DJ Katy Beauty	44	44	0	6
Eyeliner and smile Andi	10	10	0	1
Fanni Gardróbja ('Fanni's Wardrobe')	50	50	0	7
Freedee	0	-	-	-
IDDQD	67	64	3	1
Munkakerülő ('Lazybones')	2	2	0	0
NJoyStyle Reni	16	16	0	0
Pontehu	0	-	-	-
Star Wars Miniatures	0	-	-	-
Stílus és egyéb katasztrófák ('Style and other catastrophes')	45	**24**	20	1
Telefonközpont ('Phonecenter')	6	5	1	0
The very style	24	16	7	0
Vírusirtó blog ('Antivirus blog')	4	**3**	0	0
Average:	20	21.68	2.81	1.37

Table 5: Features of comments in the whole corpus

The corpus was also analyzed from the perspective of online etiquette. I only found 3 instances of users violating the norms: it occurred once in the "Stílus és egyéb katasztrófák" ('Style and other catastrophes') blog and once in the "Beautycrumbs" blog, in both cases the (female) author of the posts removed a comment (the deletion of comments is indicated in the comment thread). A more serious violation occurred in the male subcorpus: the author of "Appleblog" was involved in a verbal altercation with his readers because the information presented in one of his posts was incorrect. Four non-ideal comments were posted under the entry in question, but eventually the blogger apologized for presenting false information.

I believe that the comments, concerning their function, can correspond to Jakobson's model. However, the two genders utilize the functions according to different communicative strategies: male comments are chiefly referential. A typical (translated) example can be cited from the Appleblog: "*Just a bit of clarification: it wasn't the sales of the iPhone that increased by 112% in China, but Apple's revenue (compared to 2014 Q3!), but that's largely thanks to the iPhone*". On the other hand, comments on women's blogs seem to seek agreement, praise the author and accept norms. One comment on "The very style" blog reads: "*your vlog videos are great, hope to see more of them later :) and I can't stress enough how cuuuuute Lili is :)*". (Lili is the newborn daughter of the female blogger.)

I believe Mishne and Glance's approach can be applied to the comments in the female subcorpus, which states that comments can be considered implicit interpersonal relations (Mishne and Glance, 2006: 1). Miller and Shepherd (2004: 10) on the other hand view comments as manipulative utterances and vehicles of expressing personal opinion – this is a feature of the comments on blogs written by men: "Relationship development and social control are primarily external, directed outward, functions that use self-disclosure to build connections with others or to manipulate their opinions".

The extent to which the blogger engages in discussion with the users can also be an important qualitative feature of comments in this light: I have already mentioned that female bloggers at the end of their blog entries encourage users to post comments. The other possible means of keeping in touch with the readers is when the author also posts comments and engages in debates with his/her readers or agrees with them in the comments. However, this phenomenon was only observable in 8 blogs, although the fact that the data was collected in the summer could have a distorting effect on the overall picture.

4   Summary and outlook

In this paper, I presented a qualitative analysis of a corpus of 100 Hungarian language blog entries in the light of the trends of the international literature. The purpose of this was that the Hungarian literature only analyzes a restricted area of blogs.

My findings confirmed that representatives of the two genders exhibit different styles of blogging in terms of the frequency and length of posts. Furthermore, the comments and the replies given to them also indicate different communicative strategies. Naturally, a small-sample corpus linguistic analysis cannot conclusively define the most important features of Hungarian blogs. I could only answer my research questions partially.

1. The study mostly confirms the conclusions of the English literature concerning the communicative styles of the different genders. Women are usually the sole editors of their blogs and they respond to readers' comments more frequently. In the case of the female subcorpus the most important role of the posts was to share experiences and personal

narratives, while in the case of the male subcorpus the main purpose of posts was to comment on and interpret with credibility facts and official news. This can explain why the male subcorpus contains more external links which in turn point almost exclusively to official reports and statistics. In the case of the female subcorpus, providing external links can usually be interpreted as a social action because those links mostly point to blogs of friends.

2. From a corpus linguistic point of view, men produce longer posts and use more internal and external links and are prone to rhapsodic changes in post length. There are significantly more comments by readers and responses by the blogger in the blogs of the female subcorpus. Significant differences can also be observed concerning the temporal spread of entries: the 5 posts of the female subcorpus were published over an average of 21.3 days, while the entries in the male corpus were published over 47.8 days on average.

3. We can also find differences if we analyze the comments: the male subcorpus contains 128 comments while there are 272 in the female one. Remarkable differences also emerged in the case of comments, because they are mostly of a referential nature, that is they serve the purpose of clarifying information and they also provide new pieces of information. For readers commenting on the blogs in the female subcorpus agreement, sympathy, praise and accepting the norms are of higher importance.

In my opinion, future research should examine blogs and comments posted on blogs in a unified manner**.** There are several methods available for the analysis of comments which I could not utilize due to constraints of space, but it would be worthwhile to examine the "reaction time" of comments posted after the publication of the blog entry, and also the length and spelling of the comments. Besides, based on usernames and profiles, reading habits can also be investigated and the researcher can find out how many blogs are followed by the individual readers.

In the next phase of the research the current binary system of genders should be extended to include blogs written by members of LGBTQ groups (lesbian, gay, bisexual, transgender, queer). What I consider to be the most important task is to be able to create a corpus from online material in which the typical phenomenon of "gender role-playing" can be properly analyzed. To this end – besides blogs – forum posts also have to be included in the investigation. As a side track of this research area I consider the inclusion of posts and other "written-spoken" (Bódi, 2010a) material produced online by members of the LGBTQ community vital. Due to lack of data it was not possible to carry out such an analysis in this paper.

The relationship of blogs and social media is an important and still rather murky area. As I have alluded to it in my paper, the blog in its current form has become conventionalized, textuality is no longer an exclusive criterion and multimedia content can only be linked in blogs. In 2015 Susan Herring was discussing the features of multimodal platforms as "text plus one other mode" (Herring, 2015). However, I believe that the role of blogs has not yet crystallized completely: it would be interesting to expand the analysis to questions such as "will social media take over the role of blogs?", "what are the advantages and disadvantages of these two formats?" and "what effect will they have on each other?" These questions are interesting – as I have pointed out earlier – because most bloggers are available on a number of social media websites besides their blog.

Although social media sites can be incredibly fast channels of communication and stories can reach a vast audience, they cannot be considered safe. In the Hungarian context the takedown of a Facebook site called *Tolvajkergetők* (Thiefbusters) attracted attention in the

traditional national media[16]. 180 thousand followers were affected by the takedown of the Hungarian non-profit crime fighting organization's page, and political reasons were thought to have been the motives behind the deletion. This issue raises the question of analyzing blogs as tools of politics.

Contents of the corpus

At the time of data collection (August 6, 2016) all of the following blogs were online.

Female corpus:
http://www.theverystyleblog.hu/
http://djkatybeauty.blogspot.hu/
http://fannigardrobja.blogspot.hu/
http://stilusesegyeb.blogspot.hu/
http://njoythestyle.blogspot.hu/
http://csipkelany.blogspot.hu/
http://beautycrumbs.blogspot.hu/
http://www.breakfastatfruh.hu/
http://eyelinerandsmileandi.blogspot.hu/
http://azrail5.blogspot.hu/

Male corpus:
http://telefonkozpont.blog.hu/
http://audioworld.blog.hu/
http://swmini.blog.hu/
http://appleblog.blog.hu/
http://freedee.blog.hu/
http://munkakerulo.blog.hu/
http://iddqd.blog.hu/
http://avgnhunsub.blog.hu/
http://virusirto.blog.hu/
http://pontehu.blog.hu/

Acknowledgements

---

[16] For example, information on this topic can be found here: http://444.hu/2014/06/04/letorolte-a-facebook-a-tolvajkergetok-oldalat/ and here: http://www.kisalfold.hu/belfold_hirek/toroltek_a_tolvajkergetok_facebook_oldalat_-_tomcat_a_youtube-n_uzent/2384814/ (Accessed: 2016-09-15).

References

Bayer, J. (2005): *A háló szabadsága* ['Freedom of the net']. Budapest: Új Mandátum Könyvkiadó & Zsigmond Király Főiskola.

Blood, R. (2002): *The Weblog Handbook: Practical Advice on Creating and Maintaining Your Blog*. Cambridge MA: Perseus Publishing.

Bódi, Z. (2010a): A blogok nyelvészeti aspektusai ['The linguistic aspects of blogs']. *E-nyelv Magazin* 2010(3) http://e-nyelvmagazin.hu/2010/09/10/a-blogok-nyelveszeti-aspektusai/ (accessed 17 August 2016).

Bódi, Z. (2010b): *Infoszótár* ['IT dictionary']. Budapest: Tinta Kiadó.

Bőgel, Gy. (2006): *Blogvilág* ['Blogworld']. Budapest: HVG Kiadó.

Crystal, D. (2006): *Language and the Internet*. New York: Cambridge University Press.

Crystal, D. (2011): *Internet Linguistics – A student guide*. London & New York: Routledge.

Csala, B. (2005): Van-e olyan műfaj, hogy blog? ['Is blog a genre?'] In: Balázs, G. & Bódi, Z. (eds.): *Az internetkorszak kommunikációja* ['*The Communication of the Internet era*'], 89-110. Budapest: Gondolat Kiadó.

Csala, B. (2010): A blog műfajáról [On the genre of blog]. *E-nyelv Magazin* 2010(3). http://e-nyelvmagazin.hu/2010/09/10/a-blog-mufajarol. (accessed: 17 August 2016).

Delwiche, A. (2005): Agenda-setting, opinion leadership, and the world of web logs. *First Monday* 10(12). (5 December 2005).

Erőss, Cs. (2009): *Blogok marketing alkalmazása Magyarországon* ['The Hungarian Application of blogs in marketing']. (Unpublished MA thesis). Budapest: Budapesti Gazdasági Főiskola, Budapest, Hungary.

Gruhl, Daniel, Guha, Ramanathan**,** David Liben-Nowell, & Andrew Tomkins (2004): *Information diffusion through Blogspace*. WWW2004 blog workshop. http://www.www2004.org/proceedings/docs/1p491.pdf. (accessed: 17 August 2016).

Fishman, P. (1980): Interactional shitwork. *Heresies* 2, 99-101.

Hähn, J. (2011): *Promóciós műfajok a virtuális térben* ['Promotional genres in cyberspace']. Pécs: PTE Nyelvtudományi Doktori Iskola.

Hähn, J., Borgulya Á. (2013): The role of blogging in corporate communication and its application by Hungarian companies. In: Schmidt, M. C., Dimant, A., Nielsen, J.L.M. (eds.): *Kulturspezifik in der europäischen Wirtschaftskommunikation* 181-194. Baden: Springer-Verlag.

Hambridge, S. (1995): *Netiquette Guidelines*. https://www.ietf.org/rfc/rfc1855.txt. (accessed: 17 August 2016).

Herring, S. (1994): *Gender differences in computer-mediated communication: bringing familiar baggage to the new frontier*. http://cpsr.org/prevsite/cpsr/gender/herring.txt/. (accessed: 28 April 2016).

Herring, S.C., Scheidt, L.A., Bonus, S. & Wright, E. (2004): *Bridging the gap: A genre analysis of weblogs*. Los Alamitos, CA: IEEE. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.160.939&rep=rep1&type=pdf. (accessed: 17 August 2016).

Herring, S. (2015): *New Frontiers in interactive multimodal communication*. In: A. Georgopoulou & T. Spilloti (eds.): The Routledge handbook of language and digital communication. London: Routledge. (accessed: 17 August 2016).

Horváth, J. (2009): Hungarian university students' blogs in EFL: Shaping language and social connections. *TESL-EJ: Teaching English as a Second or Foreign Language* 12(4), 1-9.

Huszár, Á. (1994): Az együttérzés beszédaktusai [Speech acts of sympathy]. *Studia Nova*, I., 99-109.

Jóri, A. (2012): „Hiányolom a szeparált highpass filtert az envelope controller filter deviceból". Az elektronikus szaknyelvről dióhéjban ["I miss the separated high pass filter from the envelope controlled filter device". On the terminology of electronic music in a nutshell]. In: Gecső, T. & Sárdi, Cs. (eds.): *Nyelvi funkciók – stílus és kapcsolat* ['Language functions - style and contact']. Budapest: Tinta Kiadó, 139-145.

Jóri, A. (2015): Metaforikus nyelvi kifejezések vizsgálata a „funkkal szaggatott android-techno" színterén [Examination of metaphorical linguistic expressions on the "funk-dashed android techno" scene]. In: Folmeg, M. & Jóri, A. (eds.): *Világ és nyelv szenvedéllyel. Köszöntő kötet Gecső Tamás 60. születésnapjára* ['World and language with passion. Festschrift for Tamás Gecső on the occasion of his 60th birthday'], 288-296. Budapest: Tinta Kiadó.

Juhász, V. (2013). A blog műfaji attribútumai ['The attributes of the blog genre']. In: Tóth, Sz. (ed.): *Társadalmi változások – nyelvi változások. Alkalmazott nyelvészeti kutatások a Kárpát-medencében* ['Social changes – linguistic changes. Applied linguistic research in the Carpathian Basin']. Budapest–Szeged: MANYE & Szegedi Egyetemi Kiadó & Juhász Gyula Felsőoktatási Kiadó, 155–158.

Krishnamurthy, S. (2002): *The multidimensionality of blog conversations: The virtual enactment of September 11*. AOIR Internet Research 3.0: Net/Work/Theory, Maastricht The Netherlands: October 13-16, 2002.

Kumar, R., Novak, J., Raghavan & Tomkins, A. (2003): *On the bursty evolution of blogspace*. http://www1.se.cuhk.edu.hk/~hcheng/seg5010/slides/bursty-www03.pdf. (accessed: 17 August 2016).

Marlow, C. (2006): *Linking without thinking: Weblogs, readership and online social capital formation*. http://cameronmarlow.com/media/marlow-2006-linking.pdf. (accessed: 17 August 2016).

Marwick, A. (2013): 'They're Really Profound Women, They're Entrepreneurs': Conceptions of Authenticity in Fashion Blogging. In: *7th International AIII Conference on Weblogs and Social Media*, July (Vol. 8).

McNeill, L. (2005): *Genre Under Construction: The Diary on the Internet*. http://www.languageatinternet.org/articles/2005/120. (accessed: 17 August 2016).

McQuarrie, E.F., Miller, J., & Phillips, B.J. (2013): The Megaphone Effect: Taste and Audience in Fashion Blogging. Journal of Consumer Research, 40(1), 136–158. https://doi.org/10.1086/669042

Miller, C.R., Dawn S. (2004): *Blogging as Social Action: A Genre Analysis of the Weblog*.

Mishne, G., Natalie G. (2006): *An Analysis of Weblog Comments*. http://leonidzhukov.net/hse/2011/seminar/papers/www2006-blogcomments.pdf. (Accessed: 17 August 2016).

Mora, E., & Rocamora, A. (2015): Letter from the Editors: Analyzing Fashion Blogs—Further Avenues for Research. *Fashion Theory: The Journal of Dress, Body & Culture* 19(2), 149–156.

Ots, M., & Abidin, C. (2015): Commercialism, audience intimacy and brand credibility in fashion blogging. In: *Communicative Democracy. Protecting, promoting and developing free speech in the digital era.* University of Gothenburg, October 6-7, 2015.

Petykó, M. (2012): A blog műfaji jellemzőinek korpuszalapú vizsgálata ['Corpus-based analysis of the characteristics of the blog']. *Magyar Nyelvőr* 136(1), 45-72.

Pihl, C. (2014): Brands, community and style – exploring linking value in fashion blogging. Journal of Fashion Marketing and Management: An International Journal, 18(1), 3-19.

Porkoláb, Á. (2015): Az internetes kommunikáció különbségei blogokban és Facebook-oldalakon ['Differences in online communication in blogs and Facebook pages']. *Argumentum,* 11. 250-264.

Tan, D.M. (2016): Towards a new cultural paradigm: fashion blogging as a metamodern practice.

Winer, D. (2002): The history of weblogs. http://oldweblogscomblog.scripting.com/historyOf Weblogs. (accessed: 17 August 2016).

Ádám Porkoláb
University of Pécs
Doctoral Program in Applied Linguistics
Ifjúság útja 6
H-7624 Pécs
adam.porkolab@gmail.com