

A LINGUISTIC STUDY OF LANGUAGE VARIETY USED ON TWITCH.TV: DESCRIPTIVE AND CORPUS-BASED APPROACHES

Jędrzej OLEJNICZAK

Department of English Studies, University of Wrocław, Wrocław, Poland

Abstract: *This paper examines the variety of language used by the community of Twitch.tv with the application of descriptive and corpus-based methods. Twitch.tv is one of the world's most visited websites for live broadcasting of video games; the unusual conditions under which the website's users engage in discussions and express themselves gave birth to a language variety which only very remotely resembles any other internet-specific way of communicating. Twitch.tv is a website accessible from all the European, South/North American and East-Asian countries and the main language used by the majority of the site's users is English. The most fascinating aspect of the site from the linguistic point of view is the possibility for all its users to write on the chat. Approximately 400 000 users are logged in at any given time and during major broadcasts more than 100 000 users are often talking simultaneously on the same chat. These factors had an immense impact on the nature and structure of the messages. The huge interlingual gaming community (aged 16-26 on average) developed a very peculiar way to comment on the broadcasts and discuss the games being played. The language used on Twitch.tv chats is abundant in neologisms and meaningful site-specific emoticons. The corpus-based study I conducted indicates that the messages are extremely context-dependent and follow certain syntactic patterns.*

Keywords: *internet community, language variety, neologisms, emoticons, chat, eSport, computer games*

1. INTRODUCTION. THE HISTORY OF CONTEMPORARY eSPORT

In this paper I introduce and discuss concepts closely related to eSport. Therefore, in order to avoid possible misreadings or misjudgments I feel that it is necessary to closely describe the current status of professional video gaming in the media. There is little to no academic research on the subject with regard to its contemporary state, while a common belief (Gumbel, 2013) has it that the significance and the degree of this phenomenon is rather low. However, the world of eSport, as Borowy (2012: 106-108) predicted in his thesis, has evolved rapidly within the last few years. The concept of eSport has been gaining in significance since early 2000s. The gaming landscape has been shaped by increasing availability of powerful computers and worldwide access to the Internet, though back then only very few expected to make a living from playing games at the competitive, almost Olympic, level. A number of major events were organized before 2000s, but those attracted a rather limited audience and focused on games which were not truly designed to be played competitively (Popper, 2013). First real gaming tournaments with large prize pools and audiences were World Cyber Games (2000, annual ever

since) and Dreamhack (2002, biannual ever since). Players would compete in games such as Starcraft or Counter-Strike but those events had little to no impact on media until late 2000s (Kim, 2007).

Those prevalent games differed from the rest because they made players face each other instead of making them compete against highly predictable artificial intelligence. Hence, they were heavily focused on teamwork, individual player skill, split-second decision making, strategy and objective control. These factors, together with complete removal of luck elements have become characteristic of the popular eSport games and are the most significant reason for their popularity nowadays. There was a visibly growing interest of players to watch the best members of their community compete against each other. The very idea that one can watch others playing computer games the same way one watches regular sports on the television has become a deciding factor for the development of eSport into such a huge phenomenon. This was also followed by increasing number of gaming tournaments per year: The Verge reports the growth from 10/year in 2000 to 696/year in 2012 (Popper, 2013).

2010s have brought two arguably biggest competitors on the eSport market: Riot Games' *League of Legends* and Valve's *DOTA2 (Defense*

of the Ancients 2), which stemmed from the idea of *MOBA* (Massive Online Battle Arena). In these games two teams, five players each, compete against each other head to head in dynamic, fast-paced skirmishes. The games are extremely intricate and very difficult to play on the top level; they require strategic planning, excellent decision making skills, team synergy and adjusting to the other team's strategies at a pace not possible in regular sports (Minotti, 2014). All these characteristics sum up to a reason why many people find eSport so exciting to watch.

2. RESEARCH SUBJECT

2.1. The phenomenon of video game streaming. The eSport suddenly became much alike any other spectator sport and the media very quickly adjusted to that. Making use of the potential of commonly available broadband Internet connections and the streaming technology, video game streaming websites were created, which allowed for live broadcasting (broadcasting and streaming are used interchangeably for the purposes of this paper as they are used to refer to the same concept but collocate differently) of video games on a global scale. Due to the fact that the streaming technology does not have high prerequisites of its own, it also allowed the regular, individual players to live broadcast their own gaming. This was facilitated by two major game streaming services opened in the early 2010s: own3d.tv and Twitch.tv, the former of which has been closed in 2013 (Popper, 2013).

Twitch.tv has developed over the years into an incredibly popular website and was bought by Amazon in September 2014 for \$970 million (Wawro, 2014). Nowadays it is the most popular streaming service for Europe, both Americas and western Asia. It allows both normal players and eSport organizations to stream any game in any language (though English is strongly preferred by the majority). Since the site provides subscription model (in which players are allowed to subscribe to a given person for \$5/month for extra privileges), those particularly popular streamers use the website as their major source of income.

2.2. The chat feature of Twitch.tv. The reasons for the growing popularity of the video game streams are many. Firstly, due to the complexity of the games being streamed, people often watch the broadcasts to learn from the best and improve. The eSport tournament broadcasts on the other hand do not differ much from other

regular sport broadcasts (in many European countries the major eSport events are nowadays available to watch on sport TV stations and in the movie theaters). Those events take place live in the regular sport venues and receive professional commentary and coverage in multiple languages.

All the above described activities are facilitated by the website's chat feature. Every broadcast has its own chat room, a large section of the screen in which messages of other broadcast viewers can be viewed and a text box through which one can enter and send their own messages. The individual streamers are usually very open to any queries and try to pass over their insights regarding the game; they also actively interact with their audience between the games via chat. Those successful individual streamers have devoted fan communities which develop their own language varieties. The eSport game broadcasts also allow the spectators to actively participate in the chat. Due to the nature of the competitive events (normally 50 000 - 500 000 users watching live), the chat's nature differs significantly from the individual streams. Due to a relatively low message uptime (see Fig. 1.) and huge number of participants, the chat is used mostly to express emotions as any kind of discussion is impaired by the constant influx of messages. However, this kind of situational context is truly unique: potentially hundreds of thousands of humans from different backgrounds engage in communication and attempt to simultaneously convey their attitudes and thoughts.

3. RESEARCH METHODOLOGY

3.1. Corpus method: description of the research material. I created two types of corpora for the purposes of this paper. Firstly, there are 21 corpora containing 17 500 words total from the tournament broadcasts. These corpora are the major focus of the research I conducted as they reflect the behaviour of the community in an unorthodox communicative context, *i.e.* an unusually large group of users is allowed to simultaneously express their thoughts and interact. These corpora are divided into two major categories depending on the stream active viewer count: first category (hereinafter referred to as 50 000 Sample) shows the chat behaviour of 50 000 - 80 000 active users and the second category (hereinafter referred to as 150 000 sample) reflects the behaviour of 150 000 active users. Secondly, I also created 10 reference corpora to describe the

user behaviour in small, individual player streams. These are divided into 1 000 active users broadcasts (hereinafter referred to as 1 000 Sample) and 10 000 active users (hereinafter referred to as 10 000 Sample) categories. It has to be noted that the data samples from the smaller corpora are less reliable; the data collection process for the individual broadcasts is much slower and the specific streams differ from one another with regard to language variations and user types. Regardless, there are certain distinctive features relevant to those types of streams.

3.2. Methods of data collection and analysis.

I manually gathered the chat records from respective streams on Twitch.tv and pre-edited the corpora to allow their further examination with computer software. I processed the corpus data with the use of WordSmith Tools 6.0 and applied Wordlist and Keyness functions together with a number of manual operations to calculate and measure the desired statistics. Due to the nature of research material, I chose to analyse the sentence (message) length and emoticon density; I also manually measured onscreen uptime of the messages.

4. RESULTS AND ANALYSIS

4.1. Active viewer count's influence on the content of the messages. The core aim of this part of the paper is to depict the relationship of viewer count with the content of the messages sent by individual users.

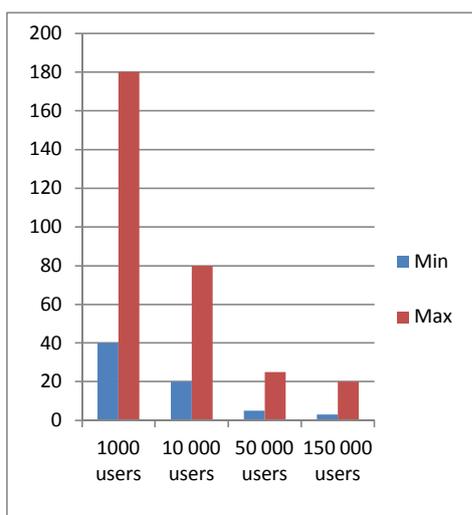


Fig.1 Message uptime (seconds)

The message uptime was also measured as I also expected it to be connected to active viewer count and thus to correlate with the message length. The purpose of this part of my research was to examine the user behaviour in unusual communicative context. The results are presented in the graphs and discussed below (see Fig. 1. and Fig. 2.).

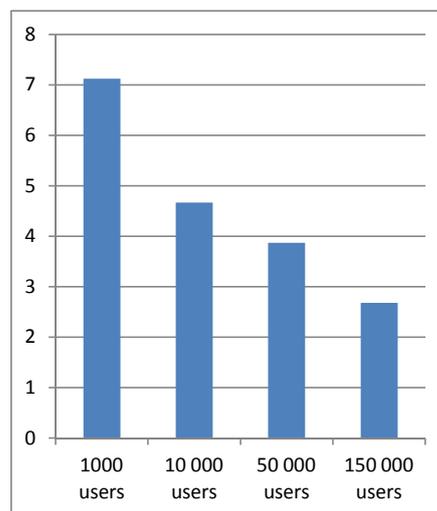


Fig.2 Sentence length (words)

The 150 000 Sample and 50 000 Sample are characterized by extremely low message uptimes (Fig.1.) and particularly short sentences (Fig. 2.). The relationship between message uptime and sentence length is intricate; the message uptime for 150 000 Sample is only slightly smaller than for 50 000 Sample, indicating that the number of messages sent does not differ drastically. The sentences, however are extremely short in 150 000 Sample and visibly longer in 50 000 Sample. Both large samples also contain unusually high amount of emoticons (see Fig. 3. and Section 4.3.). The 10 000 Sample is characterized by medium message uptime and relatively low sentence length. For the 1 000 Sample the message uptime is extremely long and the sentence length is visibly higher if compared to any other sample.

Having conducted the follow-up empirical examination of the corpora I confirmed that when communicating during the larger events (150 000 and 50 000 Samples), the users of Twitch.tv will focus on expressing themselves at all, being aware that what they write is very unlikely to be noticed and read (Fig. 1. and Fig. 2.). This means that the content of their messages will be very often limited

to two-four words, one of which is likely to be an emoticon (see Fig. 3.) and the other three will frequently contain players' names. In this case the Twitch.tv's user behaviour appears to resemble that of the crowd during major sport events. The users will cheer and even shout (write in upper case letters) knowing that there is little to no time for meaningful exchange of thoughts.

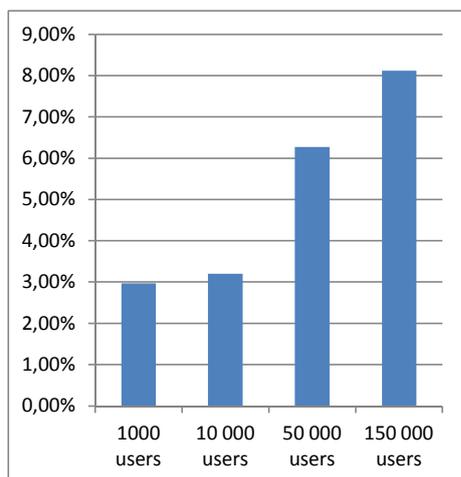


Fig.3 Emoticon % content in Twitch.tv messages

The smaller samples, unlike the large ones, are characterized by the reduced use of emoticons (more than two times less than in the large samples) and by much longer sentences. In 10 000 Sample broadcasts the users frequently engaged in exchange of thoughts, but the focus was still on relating to the gameplay (hence the medium sentence length, Fig. 2.). Longer message uptime ensured a larger variety in language and greatly limited the use of emoticons if compared to the larger samples (Fig. 1. and Fig. 2.) . The 1 000 Sample was characterized by very long message uptime, which encouraged the chat users to interact with each other and participate in meaningful exchange of opinions and thoughts. It is worth noting that the users in 1 000 Sample were often familiar with each other and their discussions sporadically encompassed topics not related to the game being played or computer games at all. Surprisingly, the use of emoticons does not differ significantly between 1 000 and 10 000 Samples (Fig. 3.), which indicates that the emoticons are used rather evenly both for commenting on gameplay and to discuss matters at hand.

4.2. The repetition principle. When I was conducting the empirical study I observed a

phenomenon closely connected to the larger samples (50 000 and 150 000). When typing in their messages, the chat users will frequently reproduce messages of other users. This results in strings of hundreds of repeated messages with exactly the same or similar content. This phenomenon usually transpires when something significant occurs during the broadcast itself. The repeated messages are usually one-two word long and contain simple phrases such as "wow" or "ok" which are sometimes accompanied by an emoticon.

The repetition principle has its variation in the smaller Samples (1000 and 10 000). In those cases, however, it takes a different shape. The community members of a given broadcast create short, nonsensical and deliberately ungrammatical stories which they later attempt to circulate. Those stories will normally end with "please do not copy-pasta *Kappa*" (which is a false request not to "copy paste", *i.e.* reproduce them). The successful ones are very quickly adopted: other users copy them and send them as their own messages. Many of those "copy-pasta" stories are later spread to other broadcasts.

4.3. The phenomenon of Twitch.tv emoticons. Contemporary research on emoticons focuses mainly on the manner of their adoption. Park *et al.* (2013, 472-473) assume that the new emoticons spread in two main ways: they are either created by a person and spread through their social networks or become spontaneously adopted. Park *et al.* (2013, 474) also hypothesize that

the emoticons are evolving from a universal way of expressing faces in text to culturally-bounded emotional dialects.

For Twitch.tv both statements hold true while the latter is of particular importance. Twitch's users use language mainly to respond to rapidly changing visual stimulus; emoticons facilitate transparent and quick expression and this is why they are often used instead of regular words (Fig. 3.). The crucial factor is that the website's owners created a feature which allows the streamers to create their own emoticons. The feature was very quickly adopted and nowadays the common emoticons have been completely replaced by the custom ones.

The custom emoticons of Twitch.tv resemble various faces, very often those of the Website's broadcasters or creators. To use the custom emoticon one has to type in a code, a simple combination of upper- and lower-case letters. The

most commonly used emoticon is encoded as "Kappa". It is a small icon resembling a grinning face cut out from the photo of one of the administrators of Twitch.tv and it denotes sarcasm and irony. It is most commonly used in the end of the messages which are intended as untrue. Interestingly, for both large Samples (150 000 and 50 000), *Kappa* was the most commonly used word and comprised approx. 2% of the all utterances (Fig. 4.). Currently, the most popular emoticons include, among others: *ResidentSleeper* (a face of the sleeping man), *PogChamp* (a screaming face), *FailFish* (face-palm expression) and *SwiftRage* (an angry face).

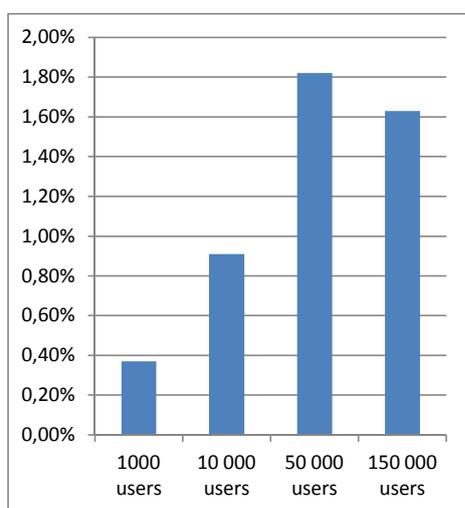


Fig. 4. Kappa emoticon % content in the text

The new emoticons are very often adopted overnight. Once added by one of the broadcasters, they may also be included by other broadcasters to their channels. This means that the broadcasters themselves control the actual spreading rate of the emoticons, while the regular users will sometimes suggest the emoticon adoption to a given broadcaster. Interestingly, this results in a situation where a given broadcaster and his/her fans frequently use their own exclusive (*i.e.* not used anywhere else) set of emoticons to communicate.

4.4. Neologisms. The community of Twitch.tv uses very peculiar means of expression, which results in constant coining of new words, some of which survive the passage of time, *i.e.* neologisation occurs (Levchenko, 2010:11-12). One of the most popular Twitch.tv neologisms, which is nowadays recognized and used by many

of the website's users has originated from the player's name. The nickname of Enrique "xPeke" Cedeño Martínez has become the part of the community's language after he singlehandedly won a seemingly unwinnable match of *League of Legends* for his team by outwitting his opponents in early 2013. The neologism "xPeke" is now recognized by UrbanDictionary. It is used in a multitude of forms, all of which appear in the corpora I investigated, including but not limited to: as a verb *to xPeke* (meaning "to outsmart"), as a noun and jokingly, as a variation of the verb "to expect" (*e.g. to xpekt*). A similar process is nowadays applied to the nickname of another *League of Legends* player, Lee "Faker" Sang-hyeok who is considered to be the best and most consistent player in the world.

Moreover, the popular Twitch.tv emoticons have become neologisms outside of the website. Even though they only have visual representations on the website itself, the expressions such as *Kappa* and *ResidentSleeper* are nowadays used by the gamer communities on various forums or in-game; the gamers assume that the meaning of those expressions is a part of the community's common ground. This also extends to the regular neologisms described above.

5. CONCLUSIONS

The conditions under which language is used on Twitch.tv do not resemble those of other internet communities; the messages are not permanent and are characterized by a low uptime on the screen, while the potentially huge number of participants imposes certain constraints on message length and content. My research revealed that the communication under those unusual circumstances favours short messages affluent in emoticons and repetitions used to express their senders' attitudes and emotions. With the decrease in the number of active users, the message length grows and regular text is increasingly favoured over emoticons. The language variety itself strives to be unique by introducing its own novel lexical elements and distinct emoticons.

Moreover, the features of language used on Twitch.tv appear to have influence on the general language of the gamer community outside of the website. The processes in which Twitch.tv neologisms diffuse into the general gamer language may prove to be an interesting research subject in the future.

BIBLIOGRAPHY

1. Borowy, M. (2012). *Public Gaming: eSport and Event Marketing in the Experience Economy*. [online]. Available: <http://summit.sfu.ca/item/12463> [Accessed on 12 April 2015].
2. Gumbel, B. (2013). *Real Sports with Bryant Gumbel: eSports Discussion*. [online]. Available: <https://www.youtube.com/watch?v=Box01YzekaQ> [Retrieved on 29 April 2015]
3. Kim, R. (2007). "League beginning for video gamers". *Sfgate.com*. [online]. Available: <http://www.sfgate.com/business/article/League-beginning-for-video-gamers-2587547.php> [Retrieved on 12 April 2015].
4. Levchenko, Y.S. (2010). *Neologism in the Lexical System of Modern English: On the Mass Media Material*. Luhansk: Luhansk Taras Shevchenko National University.
5. Minotti, M. (2014). The history of MOBAs: from mod to sensation. *Venturebeat.com*. [online]. Available: <http://venturebeat.com/2014/09/01/the-history-of-mobas-from-mod-to-sensation/> [Retrieved on 22 April 2015].
6. Park, J., Barash, V., Fink, C., Cha, M. (2013). Emoticon Style: Interpreting Differences in Emoticons Across Cultures. *Proceedings of International AAAI Conference on Weblogs and Social Media 2013*. [online]. Available: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/> [Retrieved on 25 April 2015].
7. Popper, B. (2013). Field of streams: how Twitch made video games a spectator sport. *The Verge*. [online]. Available: <http://www.theverge.com/2013/9/30/4719766/twitch-raises-20-million-esports-market-booming> [Accessed on 20 April 2015].
8. Urban Dictionary. *xPeke*. [online]. Available: urbandictionary.com/define.php?term=xpeke [Retrieved on 30 April 2015].
9. Wawro, A. (2014). Amazon to acquire Twitch. *Gamasutra*. [online]. Available: gamasutra.com/view/news/224090/Amazon_to_acquire_Twitch.php [Retrieved on 15 April 2015].