UNIVERSITY OF TARTU
DEPARTMENT OF ENGLISH STUDIES

THE ANALYSIS OF CLIPPINGS IN TWITCH LIVE STREAM CHAT ROOMS
BA thesis

HANNALEENA RÄTSEP
SUPERVISOR: *Assoc. Prof.* JANE KLAVAN, PhD

TARTU
2022

ABSTRACT

The language use of gamers and the gaming community is a relatively unexplored field of study. Due to the growing popularity of the video game streaming platform Twitch, the linguistic research of gaming language has become of even more significance. Similarly to Internet language, shortening of words is relatively common in gaming-related conversations, which raises the topic of clipping, i.e. deletion of a word part. This Bachelor's thesis discusses clippings as part of gaming language with the aim to distinguish between distinctive clipping patterns. The lack of literature regarding word formation in online communication led to a new approach of analysis which entailed the compilation of a mini-corpus using the messages of two Twitch chat rooms as the source of data.

The thesis consists of five sections: an introduction, a theoretical part, an empirical part, a conclusion, and a list of references. The theoretical part contains three sections: 1) the gradual growth of the gaming community and the need for linguistic research in the field, 2) overview of Twitch, and 3) frequent word formation processes in online communication. The empirical part presents the analysis of most frequent clippings in the messages of two Twitch chat rooms over the course of one month. The mini-corpus of the messages is accessible through a link provided in the introduction as well as in the empirical section. The thesis ends with a conclusion where the most crucial aspects of the previous section are highlighted. The list of references gives an overview of all the sources mentioned throughout the thesis.

This paper has three Appendixes. Appendix 1 illustrates the computer screen of a Twitch user, while using the chat room. Appendix 2 provides an example of a personalised Twitch home page. Appendix 3 contains twenty most frequently used clippings in the two Twitch chat rooms.

Keywords: gaming, gaming community, gaming language, Twitch, corpus, vocabulary, word formation processes, clipping

TABLE OF CONTENTS

INTRODUCTION

Today the activity of playing video games has extended beyond its enthusiasts and it has become a prominent way of entertainment across generations and social groups. Today's media provides users with a multitude of tools that can be used for better communication and allow people from all over the world to converse with each other. In the gaming community one of those tools is Twitch.tv (henceforth Twitch) − a live streaming platform that allows its users to broadcast their activities to a live chatting audience.

The gaming community was chosen as the subject of this thesis essentially due to my personal interest in video games and gaming. However, the sudden rise of players during the COVID-19 pandemic has increased the need for research in the field in general. Examining the vocabulary of the Twitch live chat would distinctly reflect the content the gamers consume and provide relevant use of gaming vocabulary. The live audience creates possibilities for various levels of discourse, which makes Twitch a significant source for investigating the language of the gaming community.

Ensslin's (2012: 7) study of "gaming language" is one of the few attempts to establish gaming jargon as a legitimate field of scholarly research. Ensslin's research introduces the concept of examining gamer interaction and communication, however she emphasises that "there is no unified language of video games and gaming" (Ensslin 2012: 7). In her research Ensslin covers various types of "gaming language", examining the differences between the language of video games ("ludological jargon"), gamer slang ("ludolect") and technical jargon ("techspeak"), and comparing their relationship to the more documented hacker language. This thesis will strictly focus on the language use of gamers, or "ludolect" as defined by Ensslin (2012: 7).

Previous studies in the gaming community have almost exclusively examined the language used in video games and the impact of it on players. Peake and Reynolds's (2020: 13) recent study revealed that in-game communication has a strong effect on the players' use of language within as well as outside the community, and supported Burwell's (2017: 42) view that video games provide an opportunity to produce new meanings. On the other hand, Strong's (2018: 274) analysis displayed an innovative use of "gamer-speak" as well as standard English between the players. Linguistic research of online gaming communities remains very limited, however.

Lack of data and limited literature will affect the chosen methods for the thesis as well as the general correlations that can be drawn between the language use of gamers and standard English. This thesis will apply one of Strong's (2018: 116) methods of collecting written data by compiling written messages of two individual Twitch chat rooms and creating a mini-corpus. The vocabulary exhibited by the users in the chat rooms will be considered "gaming language", and referred to as gaming-related or gaming-specific language.

The main focus of this study will be clippings as it was established as a distinctive linguistic feature of the two Twitch chat rooms. Clipping represents a process of shortening words and will be studied as a method of forming new words. The research question involves assessing which clipping methods occur in gaming-related conversations. Establishing Olejniczak's findings (2015: 5) as the initial foundation of the study, it is the hypothesis of this thesis that the lexical variation will be higher in the messages of a chat room with less users. The assumption is based on the belief that clippings are more common in a chat room with a higher user count, where keeping up with the chat speed is more challenging.

This thesis is divided into two chapters. The first, theoretical one, presents the growing importance of gaming, introduces Twitch and the word formation processes in the gaming community. The second, empirical one, involves description of using two Twitch chat rooms as sources of data to compile a mini-corpus of gaming-specific language (https://bit.ly/minicorpusontwitch), and analyses clippings in the messages. The thesis ends with a discussion of the results and a conclusion where the strengths and weaknesses of the thesis are given together with the description of some possible further researcher studies on this topic.

1 OVERVIEW OF THE GAMING COMMUNITY AND WORD FORMATION PROCESSES IN COMPUTER-MEDIATED COMMUNICATION

1.1 Development of the gaming community

The gaming industry has developed into one of the largest entertainment sectors in the USA and Europe. In 2021, the size of the video game market in the United States surpassed 85,86 billion U.S. dollars, more than double in comparison with 2011 when it was below 40 billion U.S. dollars (Clement 2021). The market size of the European video games industry was approximately 23,3 billion euros as of 2021. Europe has seen a yearly increase in revenue since 2017 with 2020 seeing an increase of 22% in comparison with 3% in 2019. (ISFE 2021: 18) These numbers suggest that the gaming industry has become increasingly popular during recent years.

The development of video games has certainly had an effect on the demographic of game players. A look at the annual survey statistics of The Entertainment Software Association (ESA) gives valuable information regarding the profile of adult American gamers. In 2010, 67% of the respondents played video games and the average game player was 34 years old. Precisely 75% of the players were adults with 26% being over 50 years old. (ESA 2010: 2) Similarly to 2010, the average game player was 31 years old in 2021. There has been a rise in the number of adults playing video games (80% compared to 67%), although the number of middle-aged and older players (approximately 28%) was comparable to that of 2010. (ESA 2021: 2)

A comparative study was held across European nations[1] in 2010 by GameVision Europe for the Interactive Software Federation of Europe (ISFE). According to the survey,

---

[1] There were 18 European countries included in the GameVision Europe survey: France, Germany, Italy, Spain, the UK, Portugal, Belgium, Netherlands, Switzerland, Austria, Denmark, Sweden, Norway, Finland, Czech Republic, Poland, Hungary and Latvia.

95.2 million adults had played video games in the last six months and were most popular among 16-29 year olds (57%). In contrast, there were 9% of gamers who were over 50 years old. (ISFE 2010: 15-16) In 2021 ISFE (2021: 3) joined the European Games Developer Federation (EGDF) to conduct a similar study through GameTrack and Game Sales Data (GSD)[2]. The youngest respondents were six years old, suggesting that over the years the youngest age for gamers has started to decline. Across all ages, precisely 50% played video games and the average gamer was around 31 years old. (ISFE 2021: 6)

The importance of video games during the pandemic becomes evident in the ESA survey of 2021. According to ESA (2021: 4), the gaming community helped players to find connection and sense of belonging during the pandemic. Players admitted to playing more during the pandemic (55%) as they searched for a source of relaxation as well as claimed that video games were a source of stress relief (55%) and a distraction during the difficult time (48%) (ESA 2021: 4). In Europe, video games were a source of happiness (32%), relieved anxiety (29%), helped players to feel less isolated (26%), and helped them to stay connected with their friends (26%) (ISFE 2021: 15). The stay-at-home lockdowns affected playtime, which kept fluctuating during the three years. The average playtime increased during the lockdowns and went back down when measures were lifted. (ISFE 2021: 15)

The pandemic sparked an interest towards video games; more people became involved with esports and video game streaming platforms. In fact, during COVID-19 the global audience of esports grew from 398 million in 2019 to 474 million viewers in 2021 (ISFE 2021: 19). For this reason, the Twitch gaming community appeared to be the most appropriate for the study of this thesis. At the start of the pandemic Twitch averaged about 1.27 million viewers. In April 2021, Twitch peaked in viewership with an average of

---

[2] GameTrack covered France, Germany, Italy, Spain and the UK, while GSD included data from 25 European countries.

approximately 3.1 million viewers and a maximum of about 6.25 million. Today there are approximately 2.55 million viewers on average on Twitch. There were around 3.64 million monthly broadcasters on Twitch in 2019, however today Twitch has over 8.45 million monthly broadcasters. (TwitchTracker 2021) The rise in both viewers and broadcasters demonstrates the popularity of Twitch and its possible impact on the gaming community.

1.2 Structure of Twitch

Twitch emerged as one of the live streaming platforms for video game broadcasting in 2011. The website was a branch of its mother company Justin.tv, which was created in 2007 by Justin Kan and Emmett Shear. The gaming section on Justin.tv grew exceptionally fast compared to their other live streaming categories, hence there was a need to separate the gaming content from their website. In 2014, Twitch was purchased by Amazon, although Emmett Shear remained CEO. (Stream Scheme 2021) The launch of Twitch has allowed video game live streaming to become a prominent form of entertainment. Since its emergence, Twitch has been dominating the gaming broadcasting industry. The website made an estimated 2.3 billion dollar revenue in 2020, which was primarily achieved through in-app purchases and subscriptions. (Iqbal 2022)

The users of Twitch are generally divided into two: audience and broadcasters. The broadcasters, referred to as "streamers", display their game to their audience through their computer screen − this is called video game live streaming. Twitch allows the viewers to communicate by using the live chat, also known as stream chat, next to the display, which appears on the streamer's profile page or a "channel". Every user has their own channel on which they can start live streaming at any time (Kobs et al 2019: 3). Correspondingly, every channel has their own stream chat where users can message whether the channel is live or

not (Kobs et al 2019: 3), although offline chatting is not particularly prevalent. For a complete overview of the stream chat, see Appendix 1.

In view of computer-mediated communication (CMC), Twitch chat rooms function as Internet Relay Chat (IRC), considering that Twitch is accessible to everybody on the Internet. CMC essentially refers to interpersonal communication via computer (Ferris 1997) in which conversational tools are generally categorised into two: synchronous (users talking to each other in real time) or asynchronous (users post messages and respond to them in their own time) (Papadakis 2003: 10). An average Twitch user is unable to edit or delete messages, which is often possible on platforms with asynchronous chatting, implying that Twitch chat rooms are synchronous. Romiszowski and Mason (2004: 398) further affirm that IRC is distinctive to synchronous messaging. This substantial feature of Twitch demonstrates the content transparency of the platform.

The chatting function can essentially be compared to a comment section under an online video with the exception of it being real-time and in accordance with the stream. The live chat evidently creates a community around the streamer as they are able to instantly interact with their audience (Li et al 2020: 5). The user can support the streamer by following the channel or choosing to subscribe for a monthly fee. The channels that the user follows or is subscribed to create a sidebar on their personalised home page (see Appendix 2), which always shows who is streaming at the present time. (Kobs et al 2019: 3)

Li et al (2020: 10) state that interaction between the streamer and the audience encourages the users to participate in chat conversations. Their research is supported by McMillan and Chavis's (1986: 9) concept of "sense of community", which proposes that the four elements of community are "membership, influence, integration and fulfilment of needs, and emotional connection". The users are more likely to invest their time and money

into Twitch as well as reward the streamer by subscribing to them when the interaction is frequent (Li et al 2020: 10). The ability to subscribe plays a significant part in providing the users the sense of belonging, since it divides the audience into groups. Thus, there are people who "belong" and those who do not (McMillan and Chavis 1986: 9). The latter notion may be amplified by the perks that come with a Twitch subscription, such as emotes that are frequently used in chat messages.

Emotes are emoji-like images (Kobs et al 2019: 1) characteristic to Twitch chat rooms. The users have the ability to use exclusive channel emotes added by the streamer or free ones, called global emotes, provided by Twitch. Emotes contain both stationary or moving images of people, animals and objects. Specific backstories and use cases (Kobs et al 2019: 4) of emotes make them distinctive from emoticons. Particular global emotes have a prevalent use in the Twitch community, for instance, 'LUL', 'Kappa' and 'PogChamp' (Stream Elements 2022), which express laughter, sarcasm and amazement, respectively (KnowYourMeme 2022). Aforementioned emotes indicate a type of reaction or emotion on the basis of which the purpose of emotes can be uncovered. It needs to be emphasised that every emote has a specific input, which entails correct spelling, meaning that 'PogChamp' written as 'pogchamp' will not appear as an emote.

Kobs et al (2019: 24) demonstrate a considerable use of emotes in Twitch messages in their research, suggesting that emotes have semantic components encoded in them that affect the general use of language in the chat rooms. Their study further revealed that the meaning of an emote might not be obvious without its background knowledge, and thus messages without any cue words might be challenging to analyse (Kobs et al 2019: 31). On the other hand, Ford et al (2017: 859) distinguishes "practices of coherence" that enable to make sense of the variable language in Twitch chat rooms. The study demonstrated the production of coherence through reduction of original content, which was carried out with

the help of emotes (Ford et al 2017: 865). On several occasions parts of sentences were replaced with an emote, suggesting a type of shortening or removal of content, which could be referred to as clipping (see page 26).

1.3 Word formation processes in computer-mediated communication

Chat communication is one of the most researched areas in CMC analysis (Herring 2004, as cited in Holmer 2008) given that chat discussion differs considerably from face-to-face conversations (Black et al 1983: 61). However, Ensslin (2012: 8) states that although there are various studies on the language of specific types of mass and interactive media, research of language as used in and about video games is relatively little. CMC research generally focuses on the linguistic features of chat messages, such as oral style, abbreviations and emoticons, which are considered distinctive of online language (Holmer 2008: 1). Ensslin (2012: 70) states that shortening of words to enable faster communication is characteristic to "gaming language", as well, which suggests that the standards for gaming-related vocabulary may be compared with online language conventions.

In comparison with Standard English, online language has often been modified for convenience and simplicity, although it generally follows the word formation principles of Standard English (Sun 2010: 98). The literature review will hence cover the traditional aspects of word formation, with a special emphasis on those processes that have been described as characteristic in the gaming community: acronyms and initialisms (abbreviations), affixation, blending, compounding, and clipping (Ensslin 2012: 70). The primary focus of this paper will be on clippings, however.

Kreider (1979: 5) and Bauer (1983: 237) state that acronyms form by taking the initial letters of the words in a name, title or a phrase, for instance, 'ASAP' (as soon as possible) and 'NASA'. The new word must be pronounced as a separate word not as a series of letters

in order to be considered an acronym (Bauer 1983: 237). Acronyms are generally divided into two: "letter-sounding" and "letter-recitation" acronyms (Kreider 1979: 5-6). The latter notion refers to initialisms, where the letters are pronounced separately.

Instances where acronyms are not treated as such or are pronounced letter-by-letter regardless of its structure suggest that "acronyming" is heavily based on orthography (Bauer 1983: 238). It has additionally been emphasised that acronyms are occasionally formed from letters that are not strictly the initial letters of the words in the phrase (Bauer 1983: 238). The notion is illustrated by Kreidler (1979: 7) who stresses the acronyming of the Congress of Racial Equality to CORE instead of CRE as it is not only pronounceable but also an existing word.

Schneider's study (2019: 48) demonstrates an extensive use of acronyms and initialisms in the gaming community. The illustration of the acronym 'LOL' (laughing out loud) in the gaming community and online in general shows connections between the two fields. Ensslin's (2012: 71) study supports the view that abbreviations occur frequently in gamer slang, and provides the use of initialisms in the community, such as 'XP' (experience points) and 'FPS' (first person shooter).

Affixation has been defined as a word formation process where an affix is added to the base of a word. Affixes include suffixes (following the base) and prefixes (preceding the base) (Bauer 2003: 25-27), for instance, *-er* in 'player' and *re-* in 'replay'. Affixation is generally characterised by a change in the word's part of speech, and hence could involve semantic change. For example,

(1) I like to play Valorant[3]. ('Play' is a verb.)

---

[3] Valorant is a character-based tactical first-person shooter (FPS) video game. (Valorant 2022)

(2) I am a Valorant player. ('Player' is a noun.)

Affixation often occurs with shortenings in the context of gaming. The most common suffixes in "gaming language" are *-er-*, *-ic*, *-ness* and *-ation*, as mentioned by Ensslin (2012: 71).

Blending is a process of shortening and combining two lexemes with different meanings to form a new word. Blends, also portmanteau words, entails partial loss of information of at least one of the words as illustrated by, for instance, 'frenemy' (friend+enemy) and 'emoticon' (emotion+icon). There are no apparent principles to the formation of blends, however in most cases the two words are initially shortened and then combined. (Bauer 2003: 46) Schneider's (2019: 70) research of gamer language reveals that blends appear frequently in gaming-related conversations, although there is a great distinction between those that are prevalent in Internet slang and those that appear in game vocabulary. It was found that gaming-related blends were not as frequent. (Schneider 2019: 70)

Similarly to blending, compounding involves the combining of two lexemes to create a new word. The difference between the two word formation processes is that the two compound words have no relation to their new meaning (Bauer 2003: 40), for example, 'in-box' (in+box) and 'cooldown' (cool+down). There are no particular categories of blends regarding their formation, however it has been noted that in various cases the compound word belongs under the semantic category of the root lexeme (Bauer 2003: 42). For instance, 'endgame' (end+game) refers to a type of game.

Bauer (1983: 233) defines clipping as a word formation process where a word (lexeme) is shortened without changing its meaning and word class. He stresses that frequent clipping may result in stylistic changes of a word, however (Bauer 2003: 40). According to

Bauer (1983: 233), there is no way to predict how many syllables will be clipped during the process or which word part will be deleted. An earlier definition of clippings states that the clipped word has fewer syllables than the base lexeme, however the importance of the number of syllables remains undisclosed in later descriptions. (Bauer 1983; Bauer 2003) A further exploration of contemporary studies shows that not all words that are considered clippings are shorter in syllables than their base lexeme. It can be concluded that creation of clippings is arbitrary to some degree and hence may not always form by the same conventions (Jamet 2009: 24).

Kreidler (1979: 14-21) classifies the length of clippings into two categories. The first category involves monosyllabic and polysyllabic clippings. The second category further divides polysyllabic clippings into those bearing stress on the initial syllable and those having final stress. (Kreidler 1979: 14-21) On the other hand, Bauer (1983: 233) recognizes clipping patterns based on position: final-clip, initial-clip, middle-clip and compound-clip. The terms were created for a better understanding of the position of the clippings. Removing the final part of the word (final-clip) is the most common type of clipping, while a compound-clip, where the middle is retained, is a much rarer type (Bauer 1983: 233).

Kreidler (1979: 20-21) elaborates that the clipped word is generally monosyllabic, and therefore semantically constrained. However, in some cases the clipping has retained the meaning and range of use of the base lexeme. At other times the clipped form is more restricted in context and in occurrence, which makes it difficult to determine whether the clipped forms are clippings or blends when they are used in compounds. (Kreidler 1979: 20-21) In those cases the dividing characteristic between clipped compounds and simple ones is that the former retain the compound stress (Bauer 1983: 233).

In her study Lappe (2007: 65) claims that clippings are relatively new to the linguistic scene, which supports the view of arbitrary formation of clippings. She identifies that from the 394 forms included in the combined corpus of the Oxford English Dictionary (OED) and dictionary of English slang (Beale 1989), around 4.6% of the clippings were in use before 1800. The first attestation of most clippings in the corpus (73.3%) is in the 20th century. (Lappe 2007: 65-66) Despite the recent emergence, clippings are highly characteristic to "gaming language" (Ensslin 2012; Schneider 2019). Schneider (2019: 69) claims that clippings chiefly appear with gaming-related words, such as 'crit' (critical strike) and 'frag' (fragment, fragmentation grenade, or kill), or Internet culture, such as 'fanfic' (fanfiction) and 'cosplay' (costume play).

Considering the emergence of clippings as a relatively new way of forming new words, it was determined that the aspects of clipping need further research. As clippings form a considerable part of the gaming community, the following study was conducted in chat rooms of a video game streaming platform.

2 THE LEXICAL ANALYSIS OF CLIPPINGS IN TWITCH LIVE STREAM CHATS

The empirical research of my thesis includes a corpus study carried out in two separate chat rooms on the video game streaming platform Twitch. This chapter will give an overview of the process of compiling the Twitch-based mini-corpus on gaming-specific language in chat rooms of two streamers (https://bit.ly/minicorpusontwitch), and present lexical features of the messages in the corpus. Further analysis discusses clipping as a prevalent word formation process in the chat rooms, and provides explanations for clippings.

2.1 Data collection

Prior to the data collection it was necessary to clarify the right to reduplication of user content on Twitch. A thorough review of Twitch's Terms of Service revealed that user content, such as messages, may be transmitted and stored to other devices. The corpus was created by complying with the section of User Content (§8) under Twitch's Terms of Service:

Twitch allows you to distribute streaming live and pre-recorded audio-visual works; to use services, such as chat, bulletin boards, forum postings, wiki contributions, and voice interactive services; and to participate in other activities in which you may create, post, transmit, perform, or store content, messages, text, sound, images, applications, code, or other data or materials on the Twitch Services ("User Content"). (Twitch 2021)

In order to compile the mini-corpus, two streamers were chosen for the study: the first one averaging around 150 viewers per stream, whereas the second has approximately 8,000 viewers on average as of March 2022 (TwitchTracker 2021). The names of the streamers will remain anonymous for the protection of their content and privacy of their audience. The difference in the number of viewers allows the messages to be analysed not

only on a linguistic but also on a societal level. The vaster the chat room the more pressure there is for users to keep up with the velocity of the stream. For better understanding, the streamers will be called Streamer 1 and Streamer 2, respectively. The chat rooms of respective streamers will be called Chat Room 1 and Chat Room 2.

Both streamers are variety content creators ─ the word 'content creator' is a label for someone who creates entertaining material through a medium, while 'variety' refers to the wide range of the material. The variety of video games played by the streamers ensures that their audience will use diverse vocabulary, which will allow the analysis to cover a broad area of gaming language. It is necessary to mention that both streamers occasionally go live under Just Chatting category, which is generally used when streamers want to merely talk or engage in non-gaming related entertainment with their audience with no specific game running in the background. The corpus partially covers a few Just Chatting streams, where the streamers were talking to their audience without playing a game.

The chat rooms of the two streamers have the language set on English, meaning their audience is only allowed to use English in the chat room. Monolingual chat will ensure that the vocabulary applied in the chat room and accumulated during the study will be English, although the users speaking in the chat room might be multilingual in real life. Ensslin (2012: 7) further stresses that there is no "single and unified language of video games and gaming", while Kirsimäe (2017: 10) notes in her Master's thesis, that English is commonly used to ease communication in international settings or when it proves to be the most convenient language to use. Subsequently, English often prevails as lingua franca in the gaming community. In fact, approximately 1,2 million concurrent viewers watch English speaking broadcasts on Twitch, making up about 47% of the total audience (TwitchTracker 2021).

The streams chosen for the data collection occurred over the month of March in 2022. To ensure the relevancy and unpredictability of the messages, every stream moving forward the next hour would be documented. The process of documenting took place according to the following pattern: from the start of the hour till the end of the hour every message that met the criteria for documentation would be individually copied and pasted to the corpus. In the process, the names of the users were replaced with *V* for 'viewer' and *B* for 'bot' ─ bot is an automatic chatting robot meant for moderating a chat room through previously fixated filters and commands. Any username mentioned within the messages was subtracted. It was decided that one hour will be documented in the case of Chat Room 1. Due to the larger viewer count in Chat Room 2, only half an hour was recorded each time. In total, there were eleven records of Chat Room 1 and two records of Chat Room 2.

Creating the criteria for the messages in the corpus was necessary since it was decided not to document, i.e. copy and paste, all written messages in the chat rooms. The content of a message had to meet one of the criteria in order to be included into the corpus. The primary standard was for the message to contain cue words, meaning that if the message only consisted of an emote, the message was disregarded regardless of the semantic value of the emote input. Secondly, all messages referencing games or gaming were incorporated. The rest of the criteria for the messages was created upon the understanding of most common word formation processes in the gaming community. Therefore, any message containing a form of abbreviation or shortening was included into the corpus. The previously mentioned bot messages, despite them being automated, were incorporated into the analysis when they met the criteria, as well, due to the finding that the content of bot messages influenced chat room conversations to a considerable degree.

2.2 Corpus analysis

This section discusses the general findings in the corpus, including the overall number of vocabulary items, word types, and word categories by topic. The data collected for the corpus was analysed with Laurence Anthony's corpus analysis toolkits AntWordProfiler 2.0.0 (2021) and AntConc 4.0.10 (2021). The tools were applied to the initial examination of the results as well as the primary analysis of clipping and word frequency.

A more or less equal total number of words (tokens) was chosen for the two chat rooms to ensure a comparable analysis. The number of tokens was calculated automatically by the AntConc toolkit. With a proportionate token amount in mind, the data gathered for Chat Room 1 consists of 9077 tokens while Chat Room 2 contains 9228 tokens. The similar token size of the two chat rooms allows for appropriate comparison through which it is possible to examine the lexical complexity of the chat. It was the hypothesis of this thesis that the lexical variation will be higher in the messages of Chat Room 1 with less users, considering that there will be less clippings and more information is stored in the messages. The assumption was based on the theory that the pace of conversation is characteristically faster in a chat room with more users, and hence clipping will occur more often. The lexical variation is significant to the study for it demonstrates the vocabulary size of the chat rooms. The word lists of both chat rooms illustrate their distinctive features.

The lexical complexity of the two chat rooms was determined through type-token ratio (TTR). The ratio is calculated by dividing the total number of different words (types) in a text by the total number of words (tokens). A high TTR indicates a high degree of lexical variation, whereas a low one displays the opposite. TTR is generally illustrated as a ratio, although it can be calculated into a percentage, as well. (Thomas 2005: 1) The numbers for

word types were provided automatically by AntConc. The TTR of Chat Room 1 was about 23.25% (2110/9077*100), while Chat Room 2 had a TTR of about 19.16% (1768/9228*100).

It can be concluded that the vocabulary is more varied in Chat Room 1. The calculations show that the numeral difference between the word types in the chat rooms is 342 (2110-1768). There is less lexical variety in the content of Chat Room 2, which may indicate greater use of abbreviations or clippings. The larger number of users pressures users to keep pace with the chatting which often results in omitting unnecessary information from messages as well as clipping words. From the lower lexical variety in Chat Room 2 it may be concluded that there may be a correlation between the number of users in the chat room and the diversity of vocabulary. The study has provided a significant finding that gaming vocabulary may become less flexible in larger communities.

For a complete overview of the vocabulary analysis, AntWordProfiler was used. The word list tool revealed that 24.5% of the vocabulary used in Chat Room 1 appears neither in the 2000 most frequently used English words nor among the 570 most frequently used academic words. The percentage of specialised vocabulary (35.3%) is higher in Chat Room 2, which was expected due to the larger number of tokens. In contrast, the use of vocabulary that appears among the 1000 most used English words was 68.2% and 58.4% in Chat Room 1 and Chat Room 2, respectively. The findings indicate that the vocabulary categorised as uncommon could be recognized, at least partly, as gaming-specific vocabulary.

The general data of the corpus shows a frequent use of pronouns, determiners and prepositions, which form an essential part of understanding the English language. The frequent use of essential language elements suggest that the language analysis method for the study is suitable. The five most frequently used words in the corpus altogether were 'I'

(348 uses), 'the' (322), 'a' (297), 'to' (290), and 'you' (238), which fall under the aforementioned three function parts of speech. The five most used words illustrate the credibility of the study as the top end of the frequency list of the study corresponds to the words needed for basic sentence structure. Furthermore, the data showed 1574 unique words, i.e. words that appeared only once in the corpus, which demonstrates the need for language research in the gaming community.

A closer analysis of the vocabulary used in the chat rooms reveals that there is a prevalent use of shortenings in the chat rooms. The findings suggest a comparable appearance of clippings and abbreviations in the chat rooms. Among the first 100 most frequently used words in Chat Room 1 and Chat Room 2, clippings appeared five and six times, respectively; acronyms and initialisms correspondingly occurred three and six times (see Table 1).

Table 1. The ratio between clippings and abbreviations in Chat Room 1 and Chat Room 2 among the first 100 most frequently used words

Chat Room 1		Chat Room 2	
Clippings	Abbreviations	Clippings	Abbreviations
sub	LOL	homi	YT
u	WTF	u	nt
resubbed	ASMR	sub	gg
k		merch	LOL
r		dono	DJ
		modcheck	pls

Table 1 illustrates the prevalence of shortening in the context of gaming. It can be concluded that the use of clippings and abbreviations was more frequent in Chat Room 2 due to the larger number of users. The data for Chat Room 1 strengthens the belief that clippings are

preferred in gaming-related conversations, although more vocabulary items need to be analysed for the claim.

As stated by Olejniczak (2015: 4) that emotes are a tool for responding to the "rapidly changing visual stimulus", the data in the corpus demonstrated a wide use of emotes. Inclusion of emotes into chat messages functions as evidence of the images having semantic components (Kobs et al 2019: 24). During the process of data collection, the messages raised the matter of the abundance of messages that only consisted of an emote, without any cue words. Schneider (2019: 31) stresses that users are more likely to use emotes in the chat room due to the pressure to answer quickly, which explains excessive use, or spamming, of emotes.

The corpus presents a significant characteristic of Twitch chat rooms, which entails including an emote input into the content of a message. For example,

(1) I am so happy. (No emote.)

(2) I am peepoHappy. ('peepoHappy' is an emote indicating happiness.)

The semantic foundation of emotes (Kobs et al 2019: 24) allows for formation of new words and phrases in the community. The corpus provides a variety of exemplary emotes that function as tools of showing reaction in some cases, however, have the ability to replace a word in other cases. The removal of an entire word from a message to replace it with an emote could indicate clipping to some extent.

2.3 Clippings in the corpus

This section will discuss the first 20 most frequently used clippings in Chat Room 1 and Chat Room 2 (see Appendix 3). The analysis will follow the criteria of clippings which concludes that clipped words have formed as a result of removing a part of a word and have

retained the same word class and meaning as the base lexeme. The frequency of the clippings has been provided for a better understanding of the differences in language use between the two chat rooms. In this section, initialisms CR1 and CR2 will be used for Chat Room 1 and Chat Room 2, respectively.

The ratio of syllables in the base lexeme and the clipped word is not taken into account due to the fact that a considerable amount of words in online language are already relatively short. The analysis of gaming language would not be comprehensive if a smaller number of syllables in the clipped word is considered a principal element of clippings. A great part of the study would be discarded in that case. As mentioned by Lappe (2007: 65), clippings are a rather new word formation process, which leaves room to interpret the development and criteria of clippings in context.

Final-clip is the most frequently recurring clipping type in the chat rooms. The clipping 'sub' (subscribe) was the most frequently used word in CR1 with 34 uses in comparison with the 29 times in CR2. 'Sub' appeared repeatedly among the first ten most used clippings in CR1, for instance, in 'resubbed' (17 uses) and 'subathon' (7 uses). The former indicates affixation and the addition of prefix *re-* to the shortened verb 'sub' to form 'resub' (resubscribe). On the other hand, 'subathon' is a combination of clipping and blending. Clippings in blends will be discussed in more detail later in the analysis. Additionally, CR2 exhibited the use of 'sub' as a plural noun. The suffix *-s* was added to 'sub' five separate times.

The use of final-clipping in gaming-related vocabulary is comparable to that in Internet slang words. In CR1, final-clip was more evident in gaming-related words, such as 'alt' (alternative account), 'f' (fail), 'l' (loss), 'mid' (middle), 'bot' (bottom), 'dag' (dagger), and 'diff' (different). Clippings 'bro' (brother), 'tho' (though), 'wiki' (wikipedia) and 'fuckin' (fucking) are not particularly related to games, however were used in context with

gaming in CR1. In comparison, CR2 exhibited final-clipping of gaming vocabulary through 'merch' (merchandise), 'brim' (Brimstone), 'mods' (moderator), 'valo' (Valorant), 'w' (win), and 'comms' (communication); clippings such as 'homi' (homie), 'tho' (though), 'emo' (emotional), and 'doin' (doing) were considered Internet jargon. As seen from the data, omitting the final part of a word was frequent regardless of word class.

What is of significance is that in comparison with CR1, where the most used clipping was a gaming-related term, CR2 exhibited a characteristic use of the jargon 'homi', which was the most frequently recurring clipping in CR2 with 61 uses. 'Homi' was generally added either in the front or after the name of a user who was being addressed in the chat message. A more detailed analysis establishes the prevalence of collocating 'homi' with a proper noun in CR2. For example,

(1) V: catKISS[4] chat Homi mods

(2) V: Homi [subtracted] ofc my love

In the two examples, 'homi' can be seen in use with 'chat' and 'mods' simultaneously, and a name that has been subtracted due to privacy.

The data reveals that initial-clip often appears in pronouns. The second most frequently used clipping in both chat rooms was 'u' ('you') with 33 uses in CR1 and 49 in CR2. Similarly, the use of 'ur' ('your') was rather prevalent (5 in CR1; 8 in CR2), as well. Aforementioned instances indicate clipping on the condition of pronunciation. Firstly, 'you' (/juː/) and 'u' (/juː/) have the same pronunciation considering that the letter *u* is pronounced the same as the pronoun itself. Secondly, although 'your' (/jɔː/) sounds identical with 'ur' (/jə/), the vowel of the former has a more rounded pronunciation and is articulated further back in the vowel space. The findings reveal that initial-clip is the most convenient clipping type in case of pronouns as the pronunciation of the clipped word will be similar with the

---

[4] An emote.

base lexeme. Further examination of pronouns shows that unless the pronoun gets clipped, the subject will likely be omitted from the message altogether. More instances of initial-clip in CR1 are 'k' (OK), 'cause' (because), and 'bot' (robot). The word 'bot' has two meanings in the gaming community depending on the position of the clip, hence why the word appears as an example of clipping style once more (cf. final-clipping of 'bottom' to 'bot'). Among the first 20 most used clippings in CR2, 'ello' is the only clipping besides 'u' and 'ur' where the beginning of the word has been deleted.

A distinctive type of clipping is the simultaneous deletion of both the beginning and final part of a word (compound-clip). For instance, the verb 'are' was shortened to 'r' 9 times in CR1 and 6 in CR2. The pronunciations of 'are' (/ɑː/) and 'r' (/ɑː/) are identical which may be the reason for the formation of the clipped variant. Bauer (1983: 233) emphasises that retaining the middle of the word is a rare type of clipping. Considering the limited use of aforementioned clipping in the chat rooms, the argument is plausible.

Middle-clip is a rather infrequent type of clipping. For instance, 'bday' (birthday) was used 2 times in CR1, whereas 'vibe' and 'vibes' appeared 9 and 12 times, respectively, in CR2. 'Bday' has been formed by shortening the first syllable 'birth' in 'birthday' to the initial letter of the word. It may be more challenging to distinguish 'vibe', and its plural form 'vibes', as middle-clip since the word has become a distinctive vocabulary item in informal language. 'Vibe' has formerly been studied as a clipping of 'vibraphone' or 'vibration' (Oxford English Dictionary 2022). The former suggests middle-clipping, while the latter indicates final-clipping and addition of a suffix. In CR2, the use of 'vibe' suggested it was a clipping of the verb 'vibrate', which indicates that the shortened variant has undergone middle-clipping. The second syllable of the base lexeme was shortened to just the final letter of the word. An instance, where the word has undergone clipping in the middle and final part of the word, is 'dono' (donation), which was used 18 times in CR2.

In multiple cases clippings have subsequently been combined with another lexeme. For instance, 'subathon' (7 uses in CR1) is a blend of the verb 'subscribe' and the noun 'marathon'. According to Bauer (1983: 233), it is often difficult to distinguish clippings in compounds; the difference from simple compounds is that the clipped ones receive the word stress. The stress is on the first syllable in 'subathon', hence the verb 'subscribe' has undergone clipping before blending with the noun 'marathon'. In CR2, there was a prevalent use of 'modcheck' (17 in CR2), which is a compound word of the clipping 'mod' (moderator) and the verb 'check'. What is of importance is that the word functions as an emote in CR2 and its general written form is 'modCheck', where the capital letter further distinguishes between two different words. The emote was included into the list of clippings due to its profound use as a verb, which confirms the semantic value of emotes. Examples from the corpus:

    (1) V: modCheck aint that a dagger (Use of the emote in CR1.)

    (2) V: modCheck streamer? (Emote as a verb in CR2.)

    (3) V: modCheck comms? (Emote as a verb in CR2.)

The analysis provides a significant finding regarding the use of identical clippings in CR1 and CR2. The data demonstrates that despite the considerable difference in the number of users in the chat rooms, the frequency of clippings is comparable (see Table 2).

Table 2. Frequency of clippings appearing in both Chat Room 1 and Chat Room 2[5]

	Frequency in CR1	Frequency in CR2
u	33	49
sub	34	29
r	9	6
tho	7	10
ur	5	8

[5] The presented data is not all-inclusive as it only considers 20 most frequent clippings from both chat rooms.

Similar pattern of clipping suggests that the vocabulary used in chat rooms does not depend on individual users, and there is a general understanding of the language variety in the community. It must be clarified that the scarceness of gaming vocabulary among the clippings that appear in both chat rooms is rational. During the time that the chat rooms were documented, there was no common game that was played by the streamers.

A comparable application of clipping methods can be assumed from the appearance of identical clippings in CR1 and CR2 (see Table 3) Final-clip was the most frequent clipping style in the chat rooms. The initial-clip appeared less than the final one, however proved to be The chat rooms further exhibited a limited use of compound-clip, which is considered a rare occurrence in Standard English.

Table 3. Frequency of clipping types among the 20 most frequent clippings in Chat Room 1 and Chat Room 2 respectively[6]

Clipping type	Frequency in CR1	Frequency in CR2
final-clip	14	13
initial-clip	5	3
middle-clip	1	2
compound-clip	1	2

2.4 Discussion

The clippings that appear in the corpus can be divided into two groups based on their formation: words that are a direct result of clipping, and those that have formed through clipping by another word formation process. The first group of words includes the most common clipping types, such as removing the end (final-clip), the beginning (initial-clip), or the middle (middle-clip) of the word. There were several instances where a word had undergone both the final- and initial-clip (compound-clip). There were noticeably less

---

[6] The clipping 'bot' in CR1 was considered both as a final-clip and an initial-clip due to its various meanings ('robot' or 'bottom'). Blends in both chat rooms were classified under the final-clip due to containing one.

clippings that retained the middle of the word, which corresponds to the claim that it is rather rare (Bauer 1983: 233). The second group of words were a combination of a clipping, predominantly a final-clip, and another word formation process.

The data exhibited replacement of letters and phonemes in a considerable amount of words. In numerous cases at least one letter (character representing a sound) or a phoneme (smallest unit of speech sound) in a word was switched out with a similar one, which raises the topic of changing part of a word according to its pronunciation. Removing a letter from a word to replace it with another one indicates clipping. A few instances of such word formation were 'stronk' (strong), 'bebi' (baby), 'boi' (boy), 'gurl' (girl), 'strim' (stream), and strimmer (streamer). Adhering to Bauer's (1983: 233) definition of clippings, the previous examples would not be considered clippings since the new word is the same in length with the base lexeme. However, considering the definition proposed in this thesis, the words will be acknowledged as clippings.

There is no specific name for internal modification of a word (Bauer 2003: 33). Although switching of letters as such has not been recognized as characteristic to clipping, the presented data demonstrates a word formation process that partially involves removing part of a word. Bauer (1994, as cited in Jamet 2009) additionally writes that clipping can lead to changes in spelling without there being any change in the pronunciation. Comparing the pronunciation of the clipping and the base lexeme, it can be concluded that Bauer's theory applies to words that replace a letter or a phoneme to form a new variant of itself.

Bauer (2003: 40) suggests that all things considered clipping is a way of forming lexemes rather than part of morphology. For that reason it is possible to conclude that replacement of letters in a word to form a new variant involves clipping. The data demonstrates substantial use of replacing one phoneme with another one as demonstrated above. 'Birf' (birthday), which was used once in CR1, exhibits final-clipping of the base

lexeme to 'birth', and the replacement of consonant 'th' (/θ/) with another voiceless consonant 'f' (/f/). The clipped word adheres to standard clipping rules as well as the criteria for the clippings included in the corpus analysis.

Alteration of word parts according to its pronunciation was more prevalent in CR1. The word 'easy-peasy' was shortened to 'ez pz' (used once in CR1) which indicates clipping in the middle and at the end of the word as well as evidently demonstrates the replacement of phonemes. The difference between the base lexeme and the clipping becomes more noticeable in spelling. The pronunciation of the two words is almost identical with the exception that, with the addition of 'z' (/zi:/), the vowel /i:/ in the clipping sounds longer than the initial vowel /ɪ/ in the base lexeme (cf. /i:zɪpi:zi/ vs /i:zi:pi:zi:/). The data exhibited use of 'ezpz' (3 in CR1) as 'especially', as well. The base lexeme includes two unvoiced consonants 's' and 'c' which were both replaced with the voiced 'z'. What is extraordinary is that, in comparison with numerous other clippings in the corpus, 'ezpz' is semantically rather constrained, meaning that without context it may be difficult to distinguish its word class. That explains the limited use of the clipping and further raises the question of whether the word can be classified as one at all.

On the other hand, there was a considerable use of 'ez' (easy) as it appeared 9 times in both chat rooms. The word class, and therefore the meaning, of the clipping is less challenging to guess as it has retained necessary word parts. Ensslin (2012: 73) discusses semantic shifts as a result of narrowing a meaning to a highly specific topic. It is possible that the broader use of 'ez' in the gaming community has led to the application of letter *z* in other words just as with 'ezpz' (especially).

Another unconventional use of clipping is illustrated by 'gamba' (gambling) which was used 14 times in CR2. It can be considered a clipping by Bauer's definition as well as the definition formed for the analysis in this thesis. 'Gamba' formed through final-clipping

'gambling' to 'gamb' and adding the letter *a*. Such instances where an extra letter or an affix has been added to the clipped word occurred a considerable amount of times. The claim is supported by Ensslin (2012: 71) and Schneider (2019: 49) who state that affixation comprises a part of word formation processes in the gaming community. Moreover, 'kinda' (kind of) has seemingly undergone a similar word formation process where letter *a* functions as a suffix. 'Kinda' appeared 8 times in CR1 and 13 in CR2.

Clipping of names was demonstrated by users of CR2. 'Brimstone', a character in the video game Valorant, was clipped to either 'Brim' (13 uses) or 'Brimmy' (8 uses). Although the former is an example of final-clipping, it further demonstrates the dual approach to shortenings. 'Brimmy' has been formed by adding the suffix *y* to the clipped form. Lappe (2007:) calls this type of formation "truncated names". The simple clipping as well as *y*-suffixation of names is argued to be related to the standard of addressing small children in oral speech (Schneider 2003, as cited by Lappe 2007). English nicknames often represent the referent as small for a particular communicative purpose, for example, 'Vicky' for 'Victoria' or 'Harry' for 'Harold'. Clipping of names in CR2 appeared prevalent for a similar purpose, which is to signal social belonging as well as contribute to informal communication. (Lappe 2007: 22-23)

CONCLUSION

Although studies in the fields of word formation and computer-mediated communication have been conducted by many authors in their respective fields, word formation within the gaming community is still insufficiently explored. The empirical study conducted for this thesis involved the analysis of Twitch-based mini-corpus on gaming-specific language in chat rooms of two streamers (https://bit.ly/minicorpusontwitch). Clipping was determined as one of the most prevalent word formation types in the vocabulary items used by the gaming community. Prior to further analysis, however, a revised definition was created for clipping in the context of gaming. It is evident that the impact of video games on the language use of gamers is scholastically well documented, however the literature on vocabulary in online gaming conversations is less consistent.

The formation of words in chat room conversations is relatively unique, which requires a specific approach. Hence, it was decided to modify the rules for Standard English clippings as it would allow for a more extensive analysis of vocabulary used by gamers. Examining the language as part of online jargon further provided a unique window into the implementation of vocabulary inside the community. The revised definition of clippings in the study of this thesis states that clippings are formed by removing a part of a word, while retaining the same word class and meaning as its base lexeme. Due to there being no previously fixed framework for clippings in gaming, the analysis covers a great variety of shortenings that could fall under the category; instances, where the classification is uncertain, are discussed separately.

The analysis of the mini-corpus revealed that the clipping patters in gaming-related conversations correspond to the general findings of clippings in Standard English. Final-clip was by far the most frequent type of clipping, appearing equally in gaming-related vocabulary and Internet jargon. Prevalence of final-clipping contributes to prior findings of

informal and formal language, which state that it is the most common clipping type. Middle-clip and compound-clip were not as prevalent as the initial one; however, the three were applied relatively less frequently than the final-clip. In several cases clipping occurred together with another word formation process, such as affixation, blending or compounding. The data further revealed a significant formation process of letter and phoneme replacement, which creates a possibility to examine the vocabulary items as variants of each other that are spelled or pronounced in a similar way.

The results demonstrated that clippings chiefly occurred in gaming-related vocabulary, whereas words that are generally used in Standard English usually involved other shortening processes, such as abbreviating. The data exhibited clipping in both gaming vocabulary as well as Internet jargon. The prevalence of clipped pronouns, which is characteristic to online language in general, suggests that the revised definition of clippings was reasonable in the context of the study. The detailed examination has led to the conclusion that gaming vocabulary is generally subjected to clipping, whereas words that are additionally used in Standard English usually involve other shortening processes or affixation.

It was the hypothesis of this thesis that the chat room with less users (Chat Room 1) will have a higher lexical variation than a chat room with a large number of users (Chat Room 2) due to there being more clippings in the latter's messages. The results revealed that the messages were relatively shorter in Chat Room 2 due to the rapid chat speed and possible pressure to keep up with it. The results are supported by previous research, which has concluded that a smaller number of users in the chat room results in longer messages (Olejniczak 2015: 5). However, a more detailed study is needed to conclude whether the message length is affected by clipping or by another distinctive feature of chat conversations.

This thesis addressed the use of distinctive clipping patterns in two Twitch chat rooms, however the characteristics of gaming-related conversations essentially remain unexplored. The analysis of clippings as presented in this paper could serve as a particular framework for studying clippings in an online setting. Exceptional findings signal the need for additional studies to understand more about the unique aspects of clipping, and address the question of categorising clippings based on their pronunciation and spelling. This paper ultimately illustrates that clippings form an essential part of gaming vocabulary, and strengthens the position that a more systematic and theoretical analysis is required for gaming language to be considered a legitimate field of scholarly research.

REFERENCES

Bauer, Laurie. 1983. *English Word-Formation*. Cambridge: CUP.

Bauer, Laurie. 2003. *Introducing Linguistic Morphology*. 2nd ed. Edinburgh: Edinburgh University Press.

Black, Steven, James A. Levin, Hugh Mehan and Clark N. Quinn. 1983. Real and nonreal time interaction: Unraveling multiple threads of discourse. *Discourse Processes*, 6: 1, 59-75.

Burwell, Catherine. 2017. Game Changers: Making New Meanings and New Media with Video Games. *English Journal*, 106: 6, 41-47. Available at https://library.ncte.org/journals/EJ/issues/v106-6/29180, accessed March 2022.

Clement, Jessica. 2021. Video games industry market size in the U.S. 2010-2021. Statista. Available at https://www.statista.com/statistics/246892/value-of-the-video-game-market-in-the-us/#statisticContainer, accessed April 2022.

Ensslin, Astrid. 2012. *The Language of Gaming*. Basingstoke: Palgrave Macmillan.

ESA. 2010. 2010 Essential Facts About The Computer and Video Game Industry. Available at https://www.org.id.tue.nl/IFIP-TC14/documents/ESA-Essential-Facts-2010.pdf, accessed April 2022.

ESA. 2021. 2021 Essential Facts About The Video Game Industry. Available at https://www.theesa.com/resource/2021-essential-facts-about-the-video-game-industry, accessed April 2022.

Ferris, Pixy. 1997. What is CMC? An Overview of Scholarly Definitions. *CMC Magazine*, January 1997. Available at https://www.december.com/cmc/mag/1997/jan/ferris.html, accessed April 2022.

Ford, Colin, Daniel Gardner, Leah Elaine Horgan, Calvin Liu, a. m. tsaasan, Bonnie Nardi, and Jordan Rickman. 2017. Chat Speed OP PogChamp: Practices of Coherence in Massive Twitch Chat. DOI: https://doi.org/10.1145/3027063.3052765, accessed May 2022.

Holmer, Torsten. 2008. Discourse Structure Analysis of Chat Communication. *Language@Internet*, 5: 10. Available at https://www.languageatinternet.org/articles/2008/1633, accessed April 2022.

Iqbal, Mansoor. 2022. Twitch Revenue and Usage Statistics (2022). Available at https://www.businessofapps.com/data/twitch-statistics/, accessed February 2022.

ISFE. 2010. Video Gamers in Europe 2010. Available at

https://www.isfe.eu/wp-content/uploads/2010/06/isfe_final_combined.pdf, accessed April 2022.

ISFE. 2021. 2021 Key Facts About the European Video Games Sector. Available at

https://www.isfe.eu/wp-content/uploads/2021/10/2021-ISFE-EGDF-Key-Facts-European-video-games-sector-FINAL.pdf, accessed April 2022.

Jamet, Denis. A morphophonological approach to clipping in English. OpenEdition Journals, 1: 9, 15-31. DOI: https://doi.org/10.4000/lexis.884, accessed April 2022.

Oxford English Dictionary. 2022. Available at https://www.oed.com/, accessed May 2022.

Peake, Jeni and Alexandra Reynolds. 2020. "There's a sniper on that hill!": Gaming in

English as a Global Language Environment. *OpenEdition Journals*, 11: 11, 1-26. DOI: https://doi.org/10.4000/angles.3237, accessed March 2022.

Kirsimäe, Merli. 2017. *The Compilation and Lexicogrammatical Analysis of an Estonian*

Spoken Mini-Corpus of English as a Lingua Franca. MA thesis. Department of

English Studies, University of Tartu, Tartu, Estonia.

KnowYourMeme. 2022. Available at https://knowyourmeme.com/, accessed April 2022.

Kobs, Konstantin, Albin Zehe, Armin Bernstetter, Julian Chibane, Jan Pfister, Julian

Tritscher and Andreas Hotho. 2019. Emote-Controlled: Obtaining Implicit Viewer

Feedback through Emote based Sentiment Analysis on Comments of Popular

Twitch.tv Channels. *ACM Trans. Soc. Comput*, 1:1, 1-37.

Kreidler, Charles W. 1979. Creating New Words by Shortening. *Journal of English*

Linguistics, 13: April, 24-36.

Lappe, Sabine. 2007. *English Prosodic Morphology*. Dordrecht: Springer.

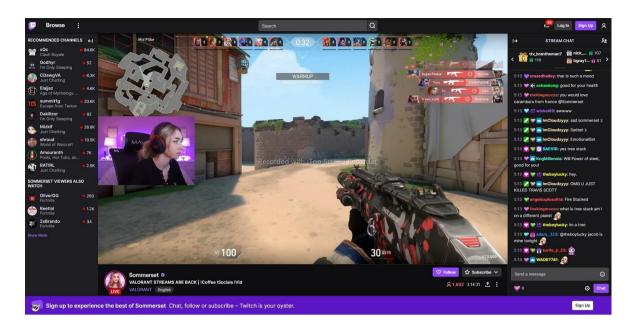Li, Yi, Chongli Wang and Jing Liu. 2020. A Systematic Review of Literature on User

Behavior in Video Game Live Streaming. *International Journal of Environmental Research and Public Health*, 17: 9, 3328.

McMillan, David W. and David Chavis. 1986. Sense of Community: A Definition and

Theory. *Journal of Community Psychology*, 14: 1, 6-23.

Olejniczak, Jędrzej. 2015. *A Linguistic Study of Language Variety Used on Twitch.Tv:*

Descriptive and Corpus-Based Approaches. Department of English Studies, University of Wrocław, Wrocław, Poland.

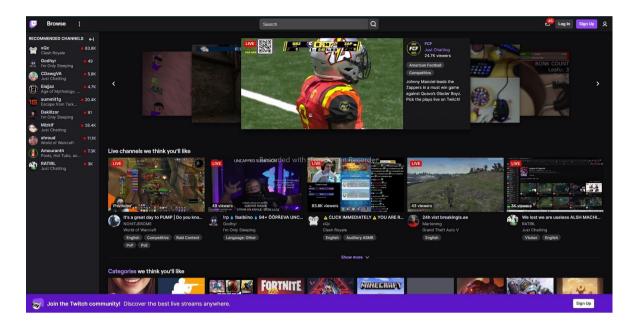Papadakis, Maria C. 2003. Computer-Mediated Communities: The Implications of

Information, Communication, and Computational Technologies for Creating Community Online. SRI International. Available at https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.207.3407&rep=rep1&type=pdf, accessed April 2022.

Romiszowski, Alexander J. and Robin Mason. 2004. Computer-mediated Communication. In David H. Jonassen (Ed.). *Handbook of Research on Educational Communications and Technology*, 397-431. New York: Routledge.

Schneider, Daniel. 2019. *English as a Lingua Franca in Online Gaming Communities and Chatrooms*. MA thesis. Department of English, University of Graz, Graz, Austria.

Stream Elements Chat Stats. 2022. Available at https://stats.streamelements.com/, accessed May 2022.

Stream Scheme. 2021. Twitch History — When Was Twitch Created? Available at https://www.streamscheme.com/twitch-a-brief-overview-and-history/#:~:text=Twitch%20History%20and%20Origins,gaming%20section%20grew%20exceptionally%20fast.&text=Twitch.tv%20was%20purchased%20by,%2C%202014%2C%20for%20%24970%20million, accessed February 2022.

Strong, Samuel. 2018. *Gamer-Generated Language and the Localisation of Massively Multiplayer Online Role-Playing Games*. PhD thesis. Centre for Translation Studies, University College London, London, England.

Sun, Hongmei. 2010. A Study of the Features of Internet English from the Linguistic Perspective. *CSCanada.net: E-Journals*, 1: 7, 98-113. Available at https://core.ac.uk/download/pdf/236303075.pdf, accessed April 2022.

Thomas, D. 2005. Type-token Ratios in One Teacher's Classroom Talk: An Investigation of Lexical Complexity. Available at https://www.birmingham.ac.uk/Documents/college-artslaw/cels/essays/languageteaching/DaxThomas2005a.pdf, accessed April 2022.

TwitchTracker. 2021. Available at https://twitchtracker.com/, accessed April 2022.

Valorant. 2022. Available at https://playvalorant.com/en-us/, accessed May 2022.

Appendix 1: Display of the live stream and the chat from the user's point of view

Appendix 2: Home page of a Twitch user

Appendix 3. 20 most frequently used clippings in Chat Room 1 and Chat Room 2

Chat Room 1		Chat Room 2	
Clipping	Frequency	Clipping	Frequency
sub	34	homi	61
u	33	u	49
resubbed	17	sub	29
k	16	merch	23
r	9	dono	18
bro	8	modcheck	17
subathon	7	brim	13
tho	7	vibes	12
alt	5	mods	11
ur	5	valo	11
wiki	4	tho	10
cause	3	ello	9
f	3	vibe	9
fuckin	3	ur	8
L	3	emo	6
mid	3	r	6
bday	2	w	6
bot	2	comms	5
dag	2	doin	5
diff	2	subs	5

RESÜMEE

TARTU ÜLIKOOL
ANGLISTIKA OSAKOND

Hannaleena Rätsep
The Analysis of Clippings in Twitch Live Stream Chat Rooms
Kärpimiste analüüs Twitchi otseülekannete jututubades
bakalaureusetöö
2022
Lehekülgede arv: 37

Annotatsioon:
      Mängijate ja mängukogukonna keelekasutus on võrdlemisi vähe uuritud valdkond. Seoses videomängude otseülekande platvormi Twitch populaarsuse kasvuga on mängukeele lingvistiline uurimine muutunud veelgi olulisemaks. Sarnaselt internetikeelele on ka mängimisega seotud vestlustes levinud sõnade lühendamine, mis tõstab esile kärpimise ehk sõnaosa eemaldamise. Käesolev bakalaureusetöö käsitleb kärpimist kui mängukeele osa, eesmärgiga eristada iseloomulikke kärpimismustreid. Puuduvad andmed sõnamoodustuse kohta veebisuhtluses tõid kaasa uue analüüsi käsitluse, mis hõlmas mini-korpuse koostamist, kasutades andmete allikana kahe Twitchi jututoa sõnumeid.

      Töö jaguneb viieks osaks: sissejuhatus, teoreetiline osa, empiiriline osa, järeldus ja kirjanduse loetelu. Teoreetiline osa sisaldab kolme peatükki: 1) mängukogukonna järkjärguline kasv ja keeleuuringute vajadus valdkonnas, 2) ülevaade Twitchist, ning 3) sagedased sõnamoodustusviisid veebisuhtluses. Empiirilises osas analüüsitakse kõige sagedasemalt kärbitud sõnu, mis ühe kuu jooksul kahes Twitchi jututoa sõnumites esinesid. Minikorpus koos sõnumitega on juurdepääsetav lingi kaudu, mis on antud nii sissejuhatuses kui ka empiirilises osas. Töö lõpeb kokkuvõttega, kus tuuakse välja eelmise osa kõige olulisemad aspektid. Kasutatud kirjanduse loetelu annab ülevaate kõigist töös mainitud allikatest.

      Sellel tööl on kolm lisa. Lisa 1 illustreerib Twitchi kasutaja arvutiekraani jututoa kasutamise ajal. Lisa 2 on näide Twitchi isikupärastatud kodulehest. Lisa 3 sisaldab kahtekümmet kõige sagedasemat kärpimist kahes Twitchi jututubades.

      Analüüsi tulemused näitasid, et mängudega seotud vestlustes esinevad kärpimismustrid vastavad üldistele järeldustele standardses inglise keeles. Sõna lõpu kärpimine oli kõige sagedamini esinev kärpimise tüüp. Sõna keskmise osa või nii esimese kui ka viimase osa eemaldamine ei olnud nii levinud kui ainult sõna esimese osa kustutamine. Neid kolme aga rakendati palju harvemini kui sõna lõpu eemaldamist. Minikorpuse andmed näitasid laiaulatuslikku tähtede asendamist sõnades, mis loob võimaluse tulevikus uurida erinevaid kärpimisi nende algse sõnavariandiga häälduse ja ortograafia põhjal.

Märksõnad: mängimine, mängukogukond, mängukeel, Twitch, korpus, sõnavara, sõnamoodustusviisid, kärpimine

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Hannaleena Rätsep,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

The Analysis of Clippings in Twitch Live Stream Chat Rooms,

mille juhendaja on Jane Klavan,

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace´i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

[allkiri]
Hannaleena Rätsep

Tartus, 24.05.2022

Autorsuse kinnitus

Kinnitan, et olen koostanud käesoleva bakalaureusetöö ise ning toonud korrektselt välja teiste autorite panuse. Töö on koostatud lähtudes Tartu Ülikooli maailma keelte ja kultuuride kolledži anglistika osakonna bakalaureusetöö nõuetest ning on kooskõlas heade akadeemiliste tavadega.

[Autori allkiri]
Hannaleena Rätsep

Tartus, 24.05.2022

Lõputöö on lubatud kaitsmisele.

[Juhendaja allkiri]
Jane Klavan

Tartus, 24.05.2022