Jihočeská univerzita v Českých Budějovicích

Pedagogická fakulta

Katedra anglistiky

Bakalářská práce

# Analysis of the Language of Internet Forums, Twitter Messages and YouTube Video Comments

Vypracoval: Ondřej Chabr, 3. ročník, ČJu–AJu

Vedoucí práce: Mgr. Jana Kozubíková Šandová, Ph.D.

České Budějovice 2017

## Declaration

I hereby declare that the thesis 'Analysis of the Language of Internet Forums, Twitter Messages and YouTube Video Comments' is my original work and that all sources used in this thesis have been duly acknowledged.

## Prohlášení

Prohlašuji, že svoji bakalářskou práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích dne 10. července 2017

………………...
Podpis studenta

**Anotace**

Tato práce se zabývá počítačem zprostředkovanou komunikací a charakteristikou jazykových prostředků pro ni typických. Analýza jazyka čerpá z korpusu sestaveného z komentářů a příspěvků na vybraných webových stránkách. Důraz je kladen na výskyt nejčastěji přítomných znaků komunikace i s ohledem na vztah k sociolingvistice. Cílem práce je předložit souhrn jazykových rysů, které jednotlivé komunity nejvíce používají, a následně porovnat komunity mezi sebou a vyzdvihnout rozdílné aspekty v jejich jazykové komunikaci.

Klíčová slova: internet, CMC, komentáře, internetová fóra, sociolingvistika, Twitter, YouTube

**Abstract**

This thesis deals with computer-mediated communication and the characteristics of language means which are typical for it. The analysis of language is based on a corpus of comments and posts from selected websites. Emphasis is placed on the occurrence of the most frequent features of communication with respect to sociolinguistics. The primary purpose of this study is to determine the language features which each community uses the most and subsequently to compare the individual communities and highlight the different aspects in their communication.

Keywords: Internet, CMC, comments, Internet forums, sociolinguistics, Twitter, YouTube

# Table of Contents

# Introduction

This bachelor thesis is concerned with an important linguistic topic in the 21$^{st}$ century – the communication of the Internet and Internet services. The Internet effected essential changes in our society. Internet communication is now spread worldwide, as new Internet users start to explore new ways to communicate and find information online. It is not a negligible value for our world because a large number of people have access to online services. This fact therefore influences not only our perception of the world but also the relationships among all potential users of the Internet.

The language we use in electronic communication has become the subject of linguistic analysis and we are all involved in its usage. Any user of Internet services may participate in creating new terminology and new possibilities to communicate quickly, easily but also accurately. Sometimes we can hear a term typical of the Internet used in a spoken conversation, which reveals how this language is now linked with our communication habits.

Popular Internet services like social networks, discussion panels or video streaming websites are places where communication among their users takes place. Though the Internet language has its specifics, each service is unique and does not have to include the same typical features. For the purpose of this thesis, I would like to pursue three popular online services and explore which features are typical for each site. As representatives of popular websites, I chose the biggest video server, *YouTube*, the platform for a quick communication, *Twitter*, and a service which has been popular since the Internet started to expand – Internet forums. Each site has its specific community of users. The theoretical part of this thesis will draw on linguistics publications, Internet sources and the concepts of sociolinguistics. I will summarise basic information about these topics and will examine in detail the sociolinguistic aspects of users who communicate online.

The aim of my thesis is to compare the three above mentioned online services with respect to the language used by their members. I will focus on the occurrence of common language features and decide how they help to achieve their users the goal of their communication. For that purpose, I need to collect a sufficiently large number of comments from these sites and analyse their structure. I will try to identify the features described in the theoretical part. I will also take into account the categories of the analysed texts and decide whether their sociolinguistic aspects correspond to the expectations. The

final results will be related to the aspects of sociolinguistics and I will consider whether the given results constitute a norm of the particular sociolinguistic group. In conclusion, all three sites will be compared and the typical expressions of their users will be defined.

# 1 The Internet

What is the Internet? A definition from *The Business Dictionary* describes the Internet as follows:

> A means of connecting a computer to any other computer anywhere in the world via dedicated routers and servers. When two computers are connected over the Internet, they can send and receive all kinds of information such as text, graphics, voice, video, and computer programs. (Business Dictionary)

In 1969, the computers of four universities in the USA were connected by the means of *ARPANET*[1]. When the *TCP/IP*[2] protocol was invented, it allowed computers to communicate among one another. In 1982, the *ARPANET* was split into a civilian network and military services. In 1990, Tim Berners-Lee invented the *World Wide Web* (WWW) at *CERN*[3] and in the second half of the 1990s, the number of people using an Internet connection exceeded 50 million (Cantoni and Tardini 28–29). Berners-Lee wrote about the Internet: "The vision I have for the Web is about anything being potentially connected with anything" (Alesso and Smith 122).

The availability of Internet access seems to be growing. As statistics from March 2017 show, almost half of the world population has got an access to the Internet. The highest percentage of users with an Internet connection is in North America – 88.1 %, while the lowest percentage is in Africa – 27.7 % (Internet World Stats). A service as widely used as the Internet is expected to have an impact on the communication among its users.

## 1.1 Internet Language

It may be problematic to speak about "Internet and language". The term "web languages" is however not appropriate for the purpose of this bachelor thesis. This term comes from another field of study and stands for coding systems which are used for constructing web pages (coding and programming languages like *HTML*, *JavaScript* etc.) (Baron 3).

---

[1] Advanced Research Projects Agency NETwork

[2] Transmission Control Protocol/Internet Protocol

[3] Conseil Européen pour la Recherche Nucléaire – European Organization for Nuclear Research

Baron (3) characterises the language of the Internet as "language issues that arise in constructing natural language to be carried across the Internet (e.g., email, chat, the content portion of Web pages)". A natural language used in daily life is not suitable for online communication and undergoes specific changes when used on the Internet. Texts on the Internet use not only basic grammatical conventions but also offer the option to use special conventions, like abbreviations or truncation of full sentences (Baron 4).

It is virtually impossible for anyone to analyse and correct the large number of emails, chat group messages or web page texts because so many new ones are added daily (Baron 4). According to Boardman (37), one day, people's ordinary websites will be important and considered as a historical document. They conserve our history in a massive way.

Regarding the similarities of Internet language and the language of speech and writing, the characteristics of the language of the Internet vary. Websites with a specific purpose are usually not so different from traditional language situations and varieties of written language can be presented on the Internet with minor stylistic changes. In the case of digital copies, there is even no change at all. Although online chats, emails or instant messages are expressed in writing, they are characterised by some speech properties. The communication requires a quick reaction (replies in chat rooms), can be deleted (as in the case of emails) or the reader can lose attention when scrolling the message off the screen. In general, this communication demands a more energetic approach and is closer to face-to-face communication (Crystal, *Internet Linguistics* 20).

Barton and Lee (164-170) wrote about comparing the strategies of the language occurring in online media with the existing modes of communication. A major question is whether communication over the Internet should be considered as speech or as writing, or possibly a hybrid of both (for more information about CMC, see chapter 2).

The body of text (comments, posts) sometimes contains other linguistics units, like emoticons or abbreviations (Holt 82). These features serve to highlight the user's individual expression and make online writing faster. At this busy time, it is necessary not only to type quickly but also to understand the principles of using these linguistics units.

## 1.1.1 Emoticons and Emoji

An important part of almost every type of informal Internet communication is the usage of emoticons or smileys. They were invented to support the online language use. Without

them, online communication is deprived not only of characteristic facial expressions and gestures of the writer but also of the writer's body posture and distance between him and others.

This limitation on expressing attitudes and opinions was the impulse to add emoticons. Every emoticon is made of a combination of keyboard characters which are typed in a sequence to show a facial expression. All characters are on a single line and often typed after a sentence, i.e., after the final punctuation mark.

Unlike emoji, emoticons are usually read sideways. There are two basic types of emoticons: one expresses a positive attitude :) and the other one a negative attitude :(. There are also versions with the nose element, i.e., :-) and :-( (Crystal, *Language and the Internet* 36).

> Emoticons have been called 'the paralanguage of the Internet' (Dery, 1993), but they are not the same, in that they have to be consciously added to a text. Their absence does not mean that the user lacks the emotion conveyed. In face-to-face communication, someone may grin over several utterances, and the effect may be noted. In Netspeak, a 'grin' emoticon might be added to just one utterance, although the speaker may continue to 'feel' the relevant emotion over several turns (Crystal, *Language and the Internet* 34).

An emoji "is a standardised set of characters" (Emojipedia) and this set is available only on the mobile and computer operating systems[4] *Windows*, *macOS*, *Android* and *iOS*. The meaning is still the same, although the graphic representation of each character can be different in each operating system. The visualisation of all emoji is presented on the website www.emojipedia.org.

In 2015, the emoji called "Face with Tears of Joy" received a special award. The award was a prestigious one because specifically this emoji was chosen as the "Word of the Year" by *The Oxford Dictionary*. It was a surprise, as this is not even a word. *The Oxford Dictionary* explained this choice by the increasing popularity of emoji on the World Wide Web in 2015. According to a research conducted by the Oxford University Press and the company *SwiftKey*, the awarded emoji constituted more than 20% of all emoji use in Britain and 17% in the USA in 2015. Emoji are very popular and their advantage is also that they facilitate communication with everyone, no matter their

---

[4] The low-level software that supports the computer's basic functions, such as scheduling tasks and controlling peripherals.

cultural and linguistic background. They transcend not only the language barrier (Danesi, *Semiotics of Emoji* 2).
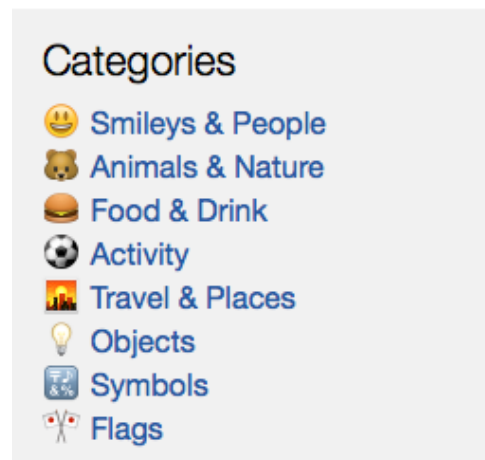


*Image 1: Categories of Emoji on Emojipedia.org*

## 1.1.2 Abbreviations

Another frequent feature of CMC are abbreviations. According to its definition, an abbreviation is a shortened form of a word or phrase. An acronym is a type of abbreviation formed from the initial letters of consecutive words and pronounced as a standard word (e.g., *ASCII* – "American Standard Code for Information Interchange", *NASA* – "National Aeronautics and Space Administration" etc.) (English Oxford Living Dictionaries). Acronyms are even used to label the technical aspects of the Internet: *HTML* – "Hypertext Markup Language", *URL* – "Uniform Resource Locator" and *DNS* – "Domain Name System" (Crystal, *Language and the Internet* 84). Another type of abbreviations are initialisms. In contrast with an acronym, in an initialism, each letter is pronounced separately and sometimes there are full stops between letters (*BBC* – "British Broadcasting Corporation") (Shortis 59).

An overview of some frequent Internet abbreviations is offered by an article by Judge Herbert B. Dixon Jr. (Dixon 30–33). Some abbreviations were well known before the Internet, such as *ASAP*, *VIP* or *TGIF*. According to Dixon, the increasing number of abbreviations is connected with the character limitation of text messaging (160 characters) and Twitter (140 characters).

# 2 CMC

The abbreviation *CMC* stands for "Computer-Mediated Communication". As the Internet gained popularity, this term became generally known in the 1990s (Crystal, *Internet Linguistics* 1). John December thinks that the definition is much more complex: "Computer-Mediated Communication is a process of human communication via computers, involving people, situated in particular contexts, engaging in processes to shape media for a variety of purposes" (December).

Different names for this type of communication are sometimes used. *EMC* stands for "Electronically Mediated Communication" and *DMC* stands for "Digitally Mediated Communication" (Crystal, *Internet Linguistics* 2).

In the past, there were coinages like *Cyberspeak* or *Netspeak*. These coinages were created for the purposes of the general public but Crystal explains that they had got a weakness: "they placed undue emphasis on the potential linguistic idiosyncrasy of the medium and suggested that the medium was more homogeneous than it actually is" (Crystal, *Internet Linguistics* 2).

Crystal (*Language and the Internet*, viii) notes that as the Internet is increasingly viewed from the social perspective, the role of the language itself becomes the main point. We can consider the Internet as a revolution and we can say that also a linguistic revolution takes place along with that.

There are many different technological variations of electronically mediated communication. The Internet brought a lot of new ways of communication. These are mostly email, chat rooms, virtual worlds, blogging, instant messaging, social networking sites, even mobile phone texting. Each of these new opportunities of communication produces a new kind of language in the case of English. "The language we use when we're blogging is not the same as the language we use when we're instant messaging" (Crystal, "How Is the Internet Changing Language Today?").

## 2.1 CMC and the Natural Language

Crystal ("How Is the Internet Changing Language Today?") also claims that the language itself has not changed that much since the CMC began to spread. For instance, we have not developed new grammar patterns and he thinks there is also not much new vocabulary. We have a few hundred new words but when we consider that there are more than a million words in English, it is not so significant. There are no changes

in pronunciation either, but we use emoticons and new punctuation has appeared. Holding a key down changes the way how users use punctuation in an exaggerated form. To emphasise the text, it is simple to use multiple exclamation marks and people have not done that before.

A few novelty features have appeared but we still deal with the same kind of English as we did before the Internet. Now we have new styles to exploit. Crystal also adds that "the language has become expressively richer as a result of the Internet" (Crystal, "How Is the Internet Changing Language Today?").

## 2.2 CMC Genres

CMC in fact created a number of new genres of communication. Brown and Attardo (187–188) give a list of the most common genres that did not exist in language before the arrival of CMC. The list contains *hypertext*, *email*, *web pages*, *listservers*, *chatting* and *blogging*.

Hypertext is used for redirecting from one text to another. This feature made the Internet popular, as it connects people and interlinks a lot of sources of information. Sometimes, parts of emails are quotes taken from the previous conversation, which provides the context. Webpages are sites where the content is presented on the Internet. They range from company business sites to personal blogs of anyone who wants to express his thoughts. Listservers create discussion groups with a central computer containing a list of email addresses. Some discussion groups are moderated (the moderator controls what is allowed to be sent and what is not), which is necessary because many off-topic messages are posted. Chatting involves usually two or more users who type their messages and are able to respond quickly. Blogging comprises online diaries but also posts about serious topics.

## 2.3 Synchronous vs. Asynchronous CMC

Each genre of Computer-Mediated Communication has its own features. The basic difference which divides the genres is that some of them consist in the type of communication where users interact at the same time and a quick response is expected. This type is called synchronous CMC. This feature is typical of online chats, where all participants are present at the same time. On the other hand, email and Internet forums are of the asynchronous type because there is often some delay between the sender's sending a message and the recipient's receiving it (Brown and Attardo 188).

# 3 Sociolinguistic Aspects Relevant to the Analysis

The behaviour of users on the Internet is as influenced by their personality as it is in real life. The question is whether the notions applied to offline interaction are only adapted to online forms of interaction or whether the form of communication changes more than that. Another question is how CMC reflects matters such as gender, race, class and ethnicity (Danesi, *Language, Society, and New Media* 27). Features occurring in posts and comments and expressions of one's own opinion are related to the above-stated social markers. That is why sociolinguistics provides important tools when examining the language of the Internet.

The human species can use language for communication, while other species cannot. The scientific discipline which examines the language and its usage is called linguistics. A more specific area of interest is sociolinguistics (Danesi, *Language, Society, and New Media* xi). People who study and explore this field are sociolinguists. Holmes (1) describes sociolinguists as professionals who "study the relationship between language and society".

Sociolinguistics began to develop in the 1960s and 1970s and dealt with a group of subjects like social class and language, languages in contact and so on. Later, computational linguistics began to develop, as in the mid-1980s personal computers became available to more users (Posteguillo 37).

According to Meyerhoff (26), it is not easy to explain all parts of this field because it is a very broad discipline. Sociolinguistics examines communication with respect to different classes and genders of speakers and also considers the situation in which the communication takes place. Everyone can change his speaking behaviour according to the situation in which he is or according to whom he is with. That is a proof of a sociolinguistic knowledge of the speaker.

## 3.1 Social Class

Every sociolinguist is "interested in how people use language and what they use it for" (Meyerhoff 27). Sociolinguists strive to find the answers to a number of questions related to society and also linguistics (Meyerhoff 28). Social classes have an impact on people's expression but it is almost impossible to determine in which social class the Internet user belongs when face-to-face interaction is missing. However, I would like to provide a brief context which I will later use for one of my analyses.

The theoretical foundation of social classes dates back to the 19<sup>th</sup> century and two persons are associated with this term – Karl Marx and Max Weber. Today, categorizing people into these classes reflects people's occupation and their wealth. Marx divided people into those who produce capital (working class) and those who control the production (middle class). The difference of Weber's point of view is that Weber included a set of behaviours, like speech, lifestyle and life chances. Social class can be defined as follows:

> A measure of status which is often based on occupation, income and wealth, but also can be measured in terms of aspirations and mobility. These factors can then be used to group individuals scoring similarly on these factors into socioeconomic classes. (Meyerhoff 156)

Determining social classes is a complex process including a lot of variables. Some sources consider people's living arrangements (higher scores are awarded to who own homes rather than those who pay rent). The most measured aspect is however one's occupation (Meyerhoff 158).

## 3.2 Politeness

Meyerhoff (84) points out that most people consider politeness as a way of expressing thoughts so as to avoid causing any offence and to show respect to others.

There are three main reasons why the Internet is less polite than face-to-face interaction. The first one is the anonymity of those who write comments. There are no consequences, except feeling guilty. The second one is the lack of bodily interaction. A face-to-face conversation requires more focus and people also use gestures. They also tend to be friendly when speaking in front of someone. The third reason concerns the delay in responding. A face-to-face conversation requires an immediate reaction (Wall).

## 3.3 Sarcasm and Irony

Sarcasm and irony are regarded as a negative speech behaviour. Usually, it is seen as hurtful and insulting but, most important, it affects the relation between the speaker and the recipient. There are almost no studies focusing on the use of sarcasm in a positive meaning. Data from the study of sarcasm by Boxer (100) prove that "sarcastic remarks

are not seen as 'slips of the tongue' but carefully planned and constructed with clear purposes in mind."

In contrast with irony, sarcasm is a type of irony but it is bitter and darker (Brown and Attardo 242). There is no agreement on whether sarcasm and irony are the same with some insignificant differences or whether they are more distinct (Salvatore 795).

# 4 Methodology

The practical part of my bachelor thesis will analyse asynchronous Internet services in terms of the language which is used for communication among members of the community. I will focus on YouTube video comments, Twitter messages (tweets) and posts from Internet forums. I have gathered a large corpus of texts from these sources and now I am going to analyse the structure of the comments to find out how many typical language features are present in each type of CMC. The selected features are typical of the language of the Internet and they are different for each source that I will examine. Each service has its specifics but I will often focus on the occurrence of abbreviations, emoticons/emoji and emphasised content. These features will be noted down and their occurrence frequency will constitute an important part in the following analysis. Besides these common features, I will focus also on several specific ones.

Sociolinguistic parameters of all types of CMC are difficult to determine when the communication takes place online, but I decided to measure what is possible and available. Categories of posts and videos will vary and each category targets a specific group of people (with exceptions), which can provide basic information about the authors of comments. The information collected will be used to analyse the language style of the different categories of members.

The last task of the practical part will be comparing the results among all of the sites. I would like to compare how the language behaviour of Twitter, YouTube and Internet forum users differs and which aspects are similar. Also, I will examine how people use profanity and impolite language when they are online and cannot be seen by others.

Detailed analyses of each video comment, tweet and Internet forum post and their language features are available on a CD enclosed to this thesis.

## 4.1 Methodology of YouTube Analysis

My language corpus is composed of comments from different categories. The most popular videos at a particular time are shown in a special section called *Popular on YouTube*. It is possible to follow this section like a classic YouTube channel. For the purpose of this thesis, I will set the language to British English, so the popular videos are showing for the area of the United Kingdom and as they were popular on 14 June 2017.

The list of videos and their categories:

- **"Super Mario Odyssey – Game Trailer – Nintendo E3 2017"** (Gaming) > 19 000 comments
- **"Simon McCoy Blooper – Drops His Career on the News! 21/1/16"** (People & Blogs) > 180 comments
- **"Grenfell Tower Resident Speaks of Social Cleansing"** (People & Blogs) > 900 comments
- **"Christian Guardino: Humble 16-Year-Old Is Awarded the Golden Buzzer – America's Got Talent 2017"** (Entertainment) > 5000 comments
- **"Olaf's Frozen Adventure – Official US Trailer"** (Film & Animation) > 3000 comments

The most recent and at the same time the most popular videos are expected to provide the newest comments and are the best source of actual commenting behaviour. I am not going to analyse the replies of other users but instead will focus on two hundred most recent comments on the first level under each video. Also, I will choose only comments in the English language to avoid the risk of losing the meaning in translation.

I will take a note whenever a part of the text or the whole text is emphasised in some way (more than three exclamation/question marks, text colour, sequence of emoticons or emoji, repeated letters or capital letters). Other features which I will record will include emoticons/emoji and abbreviations. To show how these features are distributed, I will mark down the aspect of presence of any feature.

From the sociolinguistic point of view, the comments will be grouped into four different categories. I will identify politeness and sarcasm to determine how people react after viewing the video or how they react in the comment section after reading the other posts.

## 4.2 Methodology of Twitter Analysis

Twitter is the only source in my corpus where the author of the text can be sometimes clearly identified. Still, profiles of brands or institutions can be managed by a number of editors. My analysis is based on tweets from these categories:

- Celebrity
- Brand
- Random User
- News

- Institution

As Twitter does not provide the option of a live preview with the most recent tweets, the easiest way to get access to tweets is to find user accounts and analyse tweets shown on these profiles. From the above-mentioned categories, I randomly chose the following Twitter accounts:

- **@blakeharrison23 – Blake Harrison (celebrity – British actor)**
  Registered since September 2010, 143 000 followers

- **@TeslaMotors – Tesla Motors (brand – American company)**
  Registered since February 2008, 1 480 000 followers

- **@markbakerprague – Mark Baker (random user – Journalist)**
  Registered since May 2009, 2172 followers

- **@Telegraph – The Telegraph (news – UK news site)**
  Registered since September 2008, 2 160 000 followers

- **@UniofOxford – Oxford University (institution – UK university)**
  Registered since June 2009, 367 000 followers

Given the widely different categories of users/sites selected, I expect a variety of content. I will distinguish between standard tweets, retweets and quote tweets and will note the presence of a link to another website, media content – image or video – emphasised text, emoticons/emoji, abbreviations and hashtags, which are popular and frequently used. An interesting information is the percentage of characters used up from the maximum of 140 characters available (i.e., how easy/complicated it is to express thoughts within this restriction).

Based on my own previous experience of Twitter, I am aware of what is typical for tweets and I can therefore conclude that measuring the occurrences of sarcasm/irony and politeness would not be useful. Instead, I decided to take advantage of known authors from the categories *celebrity* and *random user* and will characterise their style of writing more accurately. From each account, I will analyse the last 100 tweets before 15 June 2017.

## 4.3 Methodology of Internet Forum Analysis

Internet forums organise the posts of their users by topic. To create an appropriate corpus of posts, I will select topics with at least three replies. To obtain complex results, I decided

to analyse forums representing three different categories – entertainment, forums on a more specific theme and a discussion site about almost anything.

One of the representative entertainment forums is **GaiaOnline.com**. This site was set up in 2003 by a few comics fans. The environment of this forum is creative and entertaining. Favourite topics include movies, games or music. Teenagers can create their own space and express their personal styles. Every day, more than 300 000 users log in and spend two hours on the site. The total number of posts is over one billion (GaiaOnline). Gaia offers a huge number of topics. To avoid unfamiliar names and terms, I will skip topics focused on online games, which are very popular on GaiaOnline.
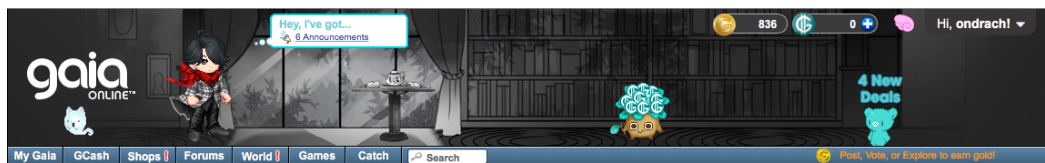


*Image 2: The Top Part of the GaiaOnline Forum*

As a forum on a more specific theme, I will analyse posts from **AutomotiveForums.com**. It was founded in 2000 and offers discussions about any car from A to Z (AutomotiveForums). I will explore topics concerning some brands, but I expect to find specialised content with many terms that are unfamiliar to me. Categories dealing with cars in general will hopefully provide me with more suitable posts for my analysis.

A discussion site on a variety of topics is represented by the popular **Reddit.com**. This site was founded in 2005 (Wikipedia). According to Alexa.com, Reddit is one of the most popular websites on the Internet (Alexa). The content is created by users who write a post, share a link or an image etc. Other users vote on the quality of the content by clicking one of the arrows, which moves the post up or down in popularity. That means the front page of Reddit is still changing and provides only the best links and other content.

There is an interesting icon indicating content which is intended for adults only. I will not work with the newest comments but will use sorting by "best" comments instead. This sorting evaluates the comments according to the number of replies and the ratio of votes indicating interesting content and votes indicating the opposite. There is also the option "sorted by top", which counts only points coming from votes by others. Other comments will be replies to the best comment.

# 5 YouTube

The Internet is full of videos and YouTube is one of the sites where videos can be uploaded. YouTube was invented by three employees of a company called *PayPal* in 2005. They came up with a new service for an easier sharing of videos and photos because email was problematic with its restrictions in formats and codecs. The beta version[5] of the new video sharing service was launched in May 2005. This version received a great feedback by users, as they viewed three million videos daily and uploaded almost eight million new videos (Sahlin and Botello 10).

Everyone can sign up for YouTube and start using the service for viewing the content of other users (their channels) or for creating their own content. To subscribe to a channel means to stay informed about new activity of the owner of the channel. In these days, the word *YouTuber* is popular among the Internet community. The definition of a Youtuber is "a person who uploads, produces, or appears in videos on the video-sharing website YouTube" (English Oxford Living Dictionaries).
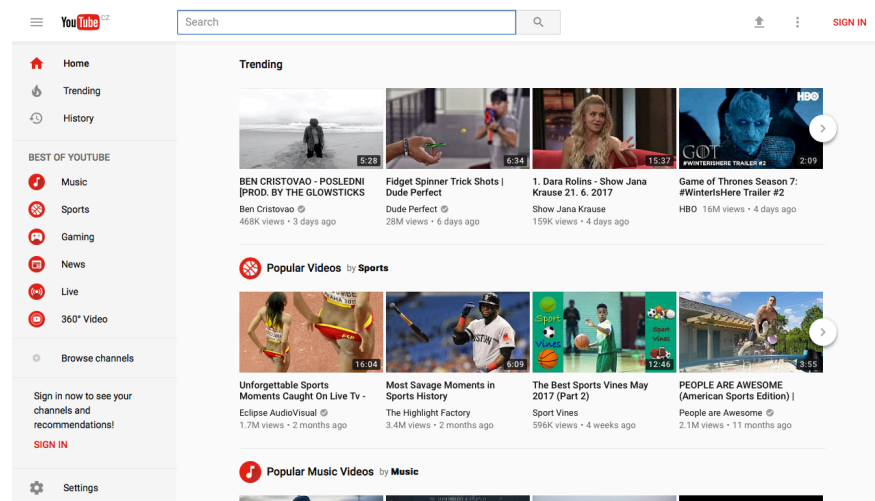


*Image 3: The Main Page of YouTube with Popular Videos*

The creators of videos are able to see detailed statistics about their videos. Some of them are visible to everyone (number of likes, views and comments) and the rest is available to the creator only (gender of viewers, age of viewers, their location etc.).

---

[5] A beta version is usually the last version before release, often tested by users under real-world conditions.

Some of the most popular videos can make their creators and actors famous. The video with the most views today is "PSY – GANGNAM STYLE"[6] with 2.8 billion views (as of June 2017). The singer Justin Bieber is now famous because of his YouTube videos (Adib).

## 5.1 YouTube Comments

All YouTube comments are public and anyone is allowed to reply. The comment section is located under the video. To prevent users from adding comments, the creator can turn comments off. Otherwise, this section is present and offers the possibility to express one's opinions or to respond to someone else's comment. Inappropriate comments can be marked as spam or abuse. A history of all comments of a specific user is also available. Comments added to videos which are deleted or comments marked as spam or abuse are also deleted (YouTube Help).

### 5.1.1 Comments Sorting

To display all comments, YouTube uses two different options: to sort by "top comments" or to sort by "newest first". The top comments option highlights comments by the author of the uploaded video, comments which generate a lively discussion or comments which are voted up by other users. The newest first option shows the most recent comments higher (according to the date and time of publication) (YouTube Help).

### 5.1.2 Voting Up

Each comment comes with the *Reply* button and also the icons of a thumb up and a thumb down. Pressing the thumb up expresses that the user likes the content of the comment, pressing the thumb down indicates that he dislikes it. This measures the popularity of each comment and is relevant to the top comments sorting. The creator can also *Pin* any comment to the top. This means that the pinned comment will remain at the top of all comments, no matter how many likes it receives or which sorting is applied.

---

[6] https://www.youtube.com/watch?v=9bZkp7q19f0

### 5.1.3 Comments Highlighting

To make their comment stand out, users can apply formatting to highlight the text. YouTube offers three possible options to change the visual appearance of the text.



- *bold text* → **bold**
- _italicized text_ → *italics*
- -strikethrough text- → ~~strikethrough~~

*Image 4: YouTube Text Formatting*

Besides these options, comments are sometimes highlighted with emoticons/emoji, capitalised text, repeated exclamation marks or question marks etc. To attract attention, commenters can use the feature of the YouTube comment section which hides text which is too long. A long comment does not necessarily need to be text, even empty new lines are accepted. In this way, the user writes something (often a question) and the rest of the comment (an answer) is hidden and visible only after clicking on *Read more*.



*Image 5: A "Read More" Comment*

# 6 YouTube Video Comments Analysis

## 6.1 Introduction to the Analysis

For the sake of clarity, I list again the analysed videos and their categories:

- **"Super Mario Odyssey – Game Trailer – Nintendo E3 2017"** (Gaming)
- **"Simon McCoy Blooper – Drops His Career on the News! 21/1/16"** (People & Blogs)
- **"Grenfell Tower Resident Speaks of Social Cleansing"** (People & Blogs)
- **"Christian Guardino: Humble 16-Year-Old Is Awarded the Golden Buzzer – America's Got Talent 2017"** (Entertainment)
- **"Olaf's Frozen Adventure – Official US Trailer"** (Film & Animation)

(Hereafter referred to by short titles, "Super Mario Odyssey", "Simon McCoy Blooper" etc.)

Videos provided a good source of texts for this study. The selection of the most popular videos in the territory of the United Kingdom provided me with the knowledge of commonly used principles of reacting to content. The comments were mostly dedicated to the theme of the video. Only a small number of them referred to a specific part in the video. These comments contained the relevant time stamp and a description of what is happening in that part of video:



*Image 6: Comment with a Time Stamp to which It Refers*

The analysed features were different for each video but the pattern of their occurrences is similar, as illustrated by the graphics that I prepared. As abbreviations, I did not count the names of products (i.e., *GTA* – Grand Theft Auto, *PS4* – PlayStation 4, *BBC* – British Broadcasting Corporation etc.) but only those abbreviations which are commonly used on the Internet to reduce the time of writing. Highlighted comments stood out immediately by repeated letters, a large number of emoticons, capital letters and other markers.

Besides highlights, emoticons and abbreviations, I noticed a relationship between the comments and the category of the video. Some of the authors of the texts were children,

while some themes were more difficult to understand, therefore the comment authors were mostly adults. I discuss this topic in more detail in chapter 6.3.

I marked two hundred most recent comments, but the video "Simon McCoy Blooper" contained only 180 comments and some of them were replies, so I decided to include the last 150 comments only.

## 6.2 Occurrences of Typical Features

After reviewing the structure of the most recent comments, I marked every occurrence of the examined characteristic. Based on the total of 950 comments on videos in the category "Popular Right Now – United Kingdom", the analysis produced the following results:
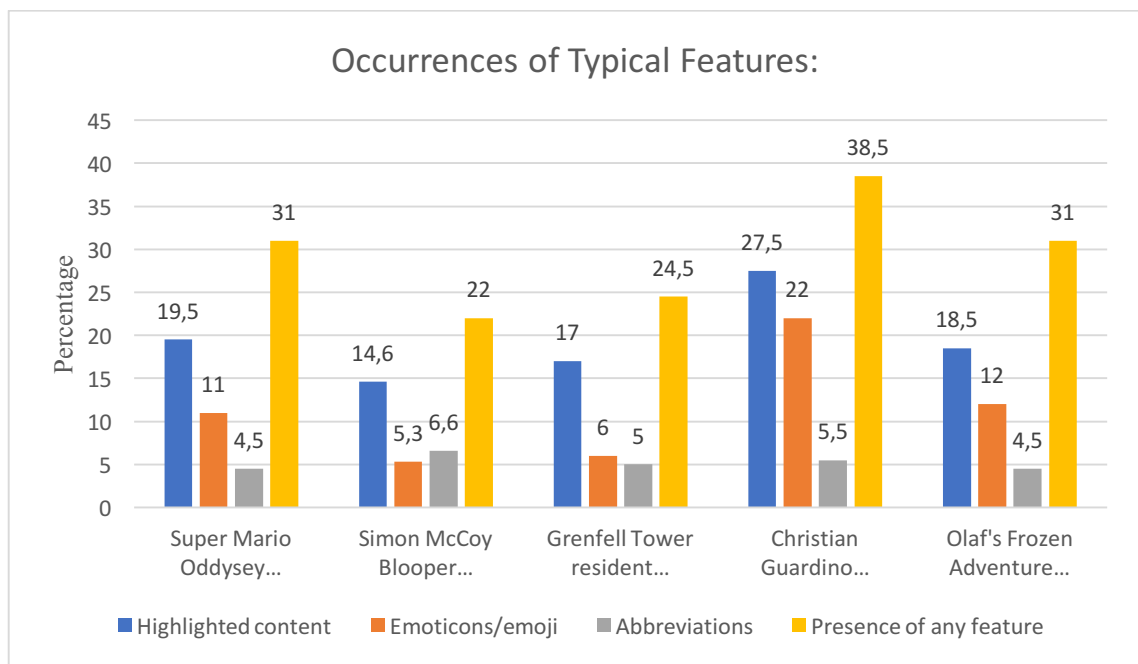


*Table 1: Occurrences of Typical Features of YouTube Comments*

Highlighted content was present more or less with the same frequency in all cases. Often by using capital letters, the users expressed their emotional response to the release of a new videogame (Super Mario) or the performance of a singer (Christian Guardino).

Regarding the use of emoticons or emoji, however, there are more obvious differences. The video about the Grenfell Tower is tragic in its content, unlike the other videos on the list, which are mostly entertaining. Accordingly, the proportion of emoticon and emoji use on this video is only 6 %. On the other hand, the video of Christian Guardino's singing generated a large amount of positive reactions and emoticons of hearts were mostly used.

The same frequency of use was evident in the case of abbreviations. Differences in percentages are minimal and the number remains almost constant. That proves that acronyms and initialism appear in a variety of categories when users want to speed up their writing.

The measure of the occurrence of any feature (emoticons, abbreviations and/or content highlighting) was always less than 50% of all comments. The rest of them were sometimes specific from the perspective of sociolinguistics.

## 6.3 Sociolinguistic Aspects

The nature of the comments is consistent with the gender and age of potential viewers from the defined category. While the category *People & Blogs* contained more sensible and longer comments, comments in the category *Entertainment* relied rather on expressing happiness and enthusiastic reactions. The minimum age limit for owning a Google account is 13 years, but it is known that even younger children have access to YouTube.

### 6.3.1 Politeness

In this section, I focused on the occurrence of any swear-words and comments with hateful references to race, religion, country etc. The video "Simon McCoy Blooper" was included in popular videos although it was published more than a year ago and that fact caused ridicule and hateful comments. Making a new short movie about the snowman Olaf was not received with enthusiasm either. The most racist comments appeared under the video of "Grenfell Tower Resident". On the other hand, the videos with a positive message about a new videogame and singing competition were well received and attracted almost no hateful comments.
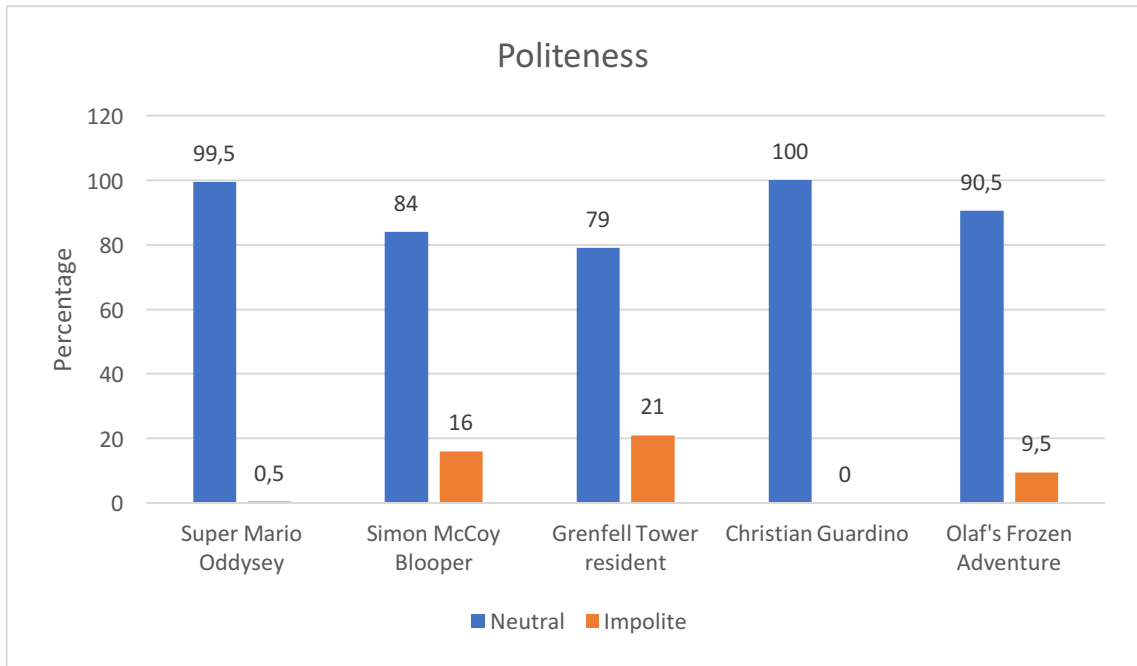
*Table 2: Politeness of YouTube Comments*

## 6.3.2 Sarcasm/Irony

As I mentioned above, the video "Simon McCoy Blooper" was highlighted as popular, although the year of its publishing is 2016. People did not realise that it is not a mistake by YouTube and apart from impolite comments, they wrote texts with ironic or sarcastic meanings, for instance:
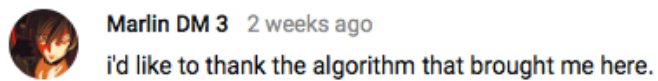


**Marlin DM 3** 2 weeks ago
i'd like to thank the algorithm that brought me here.

*Image 7: A Sarcastic YouTube Comment*

The video "Christian Guardino" was the most positive from my selection and among the last two hundred comments, there was only one which could be considered as sarcastic. "Super Mario Odyssey" was compared to other games, which was a source of sarcastic reactions. Despite a large amount of impoliteness, the most recent two hundred comments on the "Grenfell Tower Resident" video comprised only 5.5% of comments with sarcastic/ironic meaning.
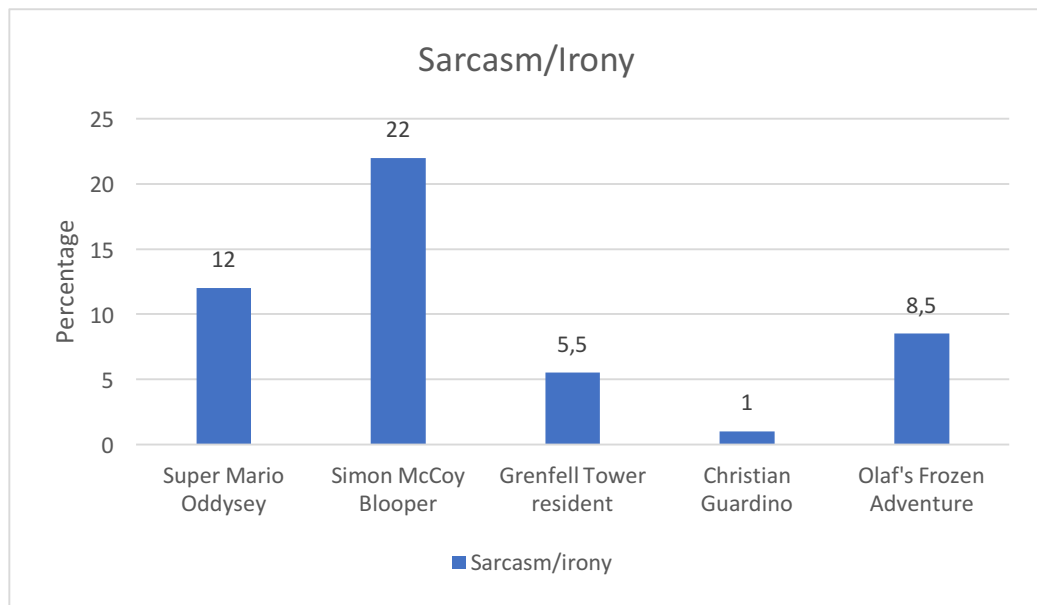
*Table 3: Sarcasm and Irony in YouTube Comments*

## 6.4 Conclusion

The expectations that I had before analysing the YouTube comments were confirmed. The nature of the comments depends on the category of video to which the comment belongs. The category *People & Blogs* guarantees a greater proportion of longer comments and a higher level of impoliteness. On the other hand, commenters in this category focused more on relevant thoughts than in the categories *Entertainment* and *Film & Animation*, where users focused on expressing their positive point of view, used a lot of emoticons and did not use much sarcasm/irony.

### 6.4.1 Examples

- Impolite comments containing swearwords (with added asterisks):

    "what the f\*\*k even happened"

    "Blah blah blah, same old sh\*t 😴"

    These comments were found under videos from the *People & Blogs* category. The authors expressed their disagreement with the topic of the video by using a swearword.

- Ironic comments exaggerate the meaning conveyed in order to express disagreement, similarly as impolite comments:

"Charterhouse student hmmm"

"I lost hope in humanity."

- Highlighted comments with emoticons appeared mostly in the *Entertainment* category:

    "He's amazing!!!😱"

    "Omg ❤❤🎵🎵"

    "💕💕amazing"

# 7 Twitter

Twitter is one of the biggest online social networks in today's Internet world. Its co-founder Jack Dorsey sent the world's first tweet on 21 March 2006. In August 2007, Chris Messina proposed the introduction of the hashtag (#) as a symbol for groups. Current statistics evidence the popularity of Twitter, there are 313 million monthly active users (Twitter).

Social networks enjoy great popularity among Internet users. The leader in the social media category is Facebook with 1.79 billion monthly active users (F Newsroom FB). In contrast with Facebook, Twitter is a specific medium which has got some other differences. In general, a Facebook user is in touch mostly with his friends, on the other hand, a Twitter user follows people because he likes them for some reason and the person does not have to be his friend.

After registering for the service on the website or in a mobile application, each user receives his own profile. It is possible to change a profile photo and the large photo in the header of the profile, i.e., background image.
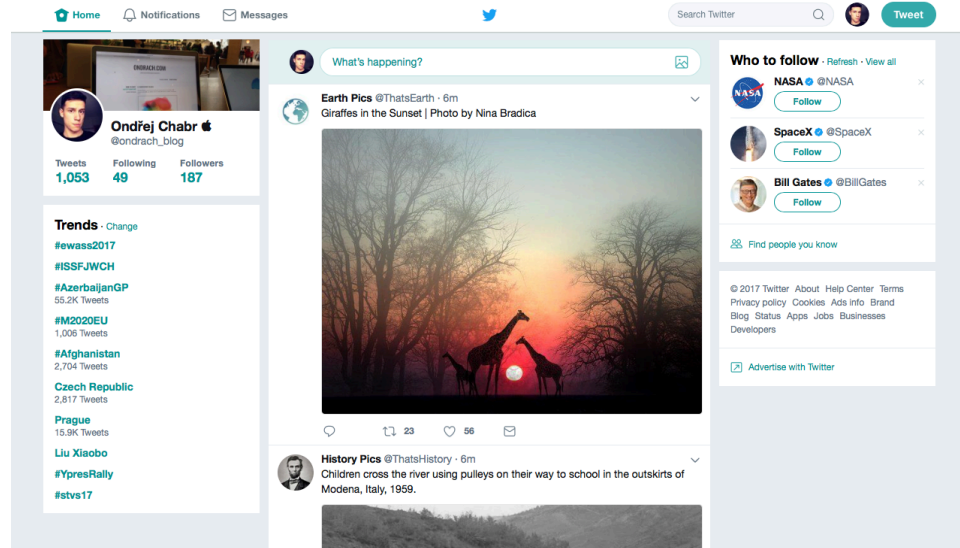


*Image 8: The Main Page of Twitter*

The main feature of Twitter is sharing texts, images, videos or animations (in the GIF format). A shared post is called a *tweet*. Every tweet is published to whomever wants to follow the user. These users are called *followers*. Posting one's own tweets is not the only way to communicate. A popular function of Twitter is *retweeting*. It consists in sharing someone else's tweet. One can either retweet someone's tweet without any

changes or add one's own comment to it. A retweet with a comment is called a *quote tweet*. Users are also able to send direct messages, which are hidden for everyone else, except the sender and the recipient.

## 7.1 Character Limitation

Ever since Twitter was launched, the policy of the service included limiting the number of characters each user can write – one post can be no longer than 140 characters. This limitation is one of the reasons why Twitter became so popular. According to *The Wall Street Journal*:

> By imposing the 140-character limit for most of Twitter's nine-year existence, the service has pressured users to write succinctly, which created the quirky conversational culture that has come to define the social-media wire. (Yoree)

The CEO[7] of Twitter, Jack Dorsey, shared a screenshot of his notes because of their length, which was impossible to contain within one tweet (Jack).

In September 2016, Twitter partly allowed to create longer tweets. Before this change, photos, videos, GIFs, polls, and quote tweets took a certain amount of space and less than 140 characters remained. After the change, however, any media content uploaded directly to Twitter, polls or quote tweets do not count towards the 140 characters (Titcomb).

The original limit on the tweet length still exists and users must be creative to shorten their tweets without any loss of meaning. Links or other sources added to a tweet are shortened but they count towards the limit.

## 7.2 The Importance of Twitter

The possibility of quick publishing and sharing media content plays an important role in today's society. In some contexts, Twitter can be more useful than traditional media and can bring the news faster than a whole article, which takes some time to write.

---

[7] Chief Executive Officer

In January 2009, a US Airways plane crashed in the Hudson River and the world knew it before journalists reported the accident. The entrepreneur Janis Krums published a photo on his Twitter account and added a brief description of the situation (Twitter).

Since then, Twitter has brought a lot of breaking news with first photos or even videos from the scene. The Boston Marathon bomb attack, the death of Whitney Houston or the Royal Wedding Announcement were also shared via Twitter before they were written by professional editors (Turner).

Everyone with a profile on Twitter and a smartphone can be the first person to spread the latest news and, in the event of some danger, for example a natural disaster, also save lives.

## 7.3 Media Content in Tweets

When a media content (image, video, GIF animation) is included in a tweet, it can influence the popularity of the tweet and generate a higher number of retweets and likes. Before October 2013, images were not visible automatically. The image opened only after clicking on the image link. Since then, every photo and image show instantly (Moritz).
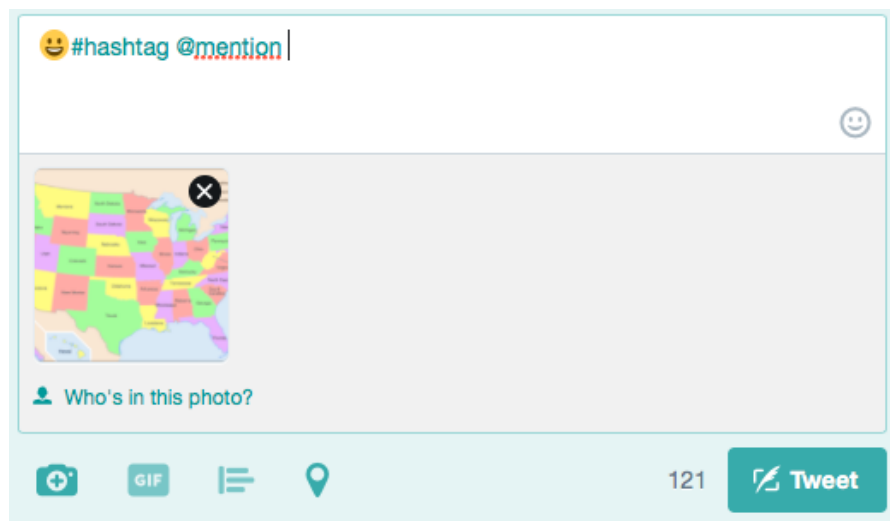


*Image 9: Composing a New Tweet with Added Features*

# 8 Twitter Messages Analysis

## 8.1 Introduction to the Analysis

The analysis of Twitter accounts contained more features typical for that service. Tweets usually include hashtags, therefore, I decided to mark down any occurrences of hashtags. Authors of tweets are restricted to 140 characters only, which sometimes led to their creativity and resulted in their making the post shorter by using an abbreviation or three dots. Some people or institutions use their accounts for sharing the content of their websites or any other content on the Internet. In that case, the posts contained a link to click on.

On Twitter, options for text formatting are missing, so to highlight a post, users resort to different tricks. Adding images and videos helps the post to attract attention or express something that is difficult or inappropriate to write. Emoticons can also be used for highlighting when more of them are placed consecutively.

These are the various kinds of accounts I chose for my analysis:

- @blakeharrison23 – Blake Harrison (celebrity – British actor)

- @TeslaMotors – Tesla Motors (brand – American company)

- @markbakerprague – Mark Baker (random user – Journalist)

- @Telegraph – The Telegraph (news – UK news site)

- @UniofOxford – Oxford University (institution – UK university)

## 8.2 Occurrences of Typical Features

My analysis is based on examining the last 50 tweets published on each account, which gives a total of 250 tweets. To start with, I marked the occurrences of their typical features.
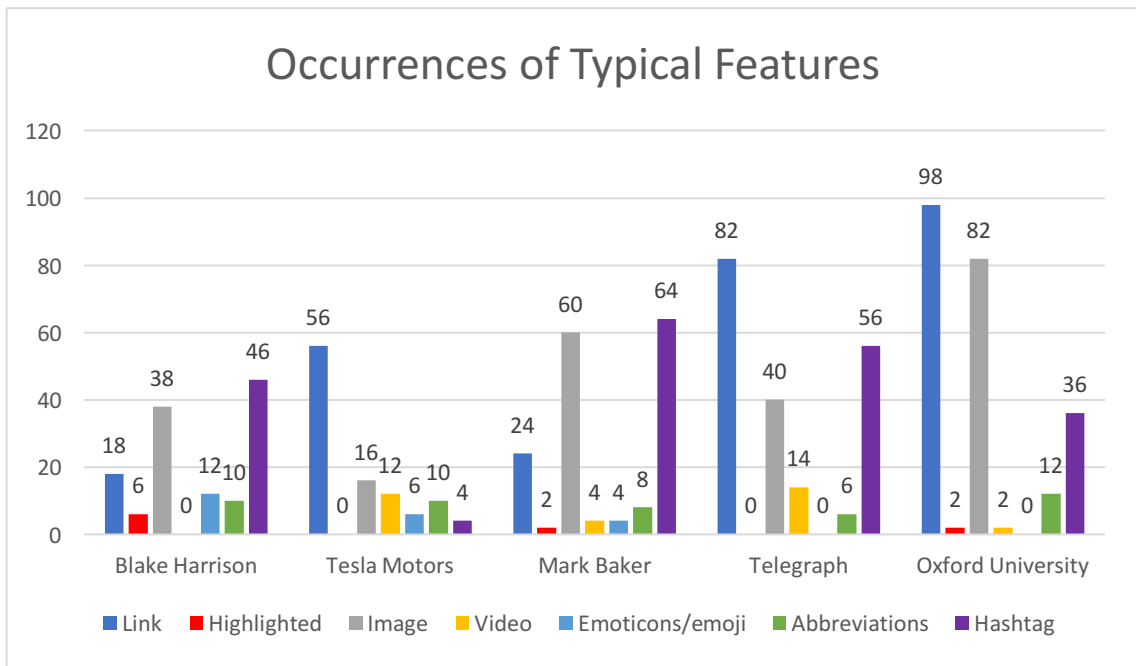
*Table 4: Occurrences of Typical Features of Tweets*

On Twitter, I found a greater number of features than on YouTube and Internet forums. The personal accounts of *Blake Harrison* and *Mark Baker* did not share links so often as other accounts, however, the tweets of *Oxford University* consisted from 98% of added links. The company brand and news accounts also made a frequent use of links.

As I already mentioned, Twitter lacks the option to highlight the text. The very low percentage of highlighted texts following from my analysis confirmed this fact. *Blake Harrison* highlighted his tweets by repeating letters or by adding more emoticons. More formal accounts did not highlight at all.

Images constitute an important part of tweets. *The Telegraph* and *Oxford University* took advantage of image content and combined the presence of an image with a link to their sites.

*Image 10: Images Attract the Attention of Visitors*

Videos in tweets are still not as widespread as YouTube videos and users prefer posting a link to YouTube rather than uploading the video directly in a tweet. Still, *The Telegraph* uploaded short videos which contained more information about breaking news for readers than an article would.

*Oxford University* and *The Telegraph* did not include any emoticons in their last 50 tweets. Other accounts did use some emoticons, but not so much. The actor *Blake Harrison* made use of emoticons to express his opinion in combination with hashtags.

The percentage of abbreviations is similar to that of emoticons. In my analysis of abbreviations in tweets, the same as in my analysis of YouTube video comments, I did not count the names of products and institutions as abbreviations. *Blake Harrison* in one of his tweets mentioned "Question Time" and also the abbreviation "QT". Surprisingly, not even *Oxford University*, *The Telegraph* and *Tesla Motors* used abbreviations.

The last feature in this part of my analysis is the hashtag. There was rarely more than one hashtag in a tweet. *Mark Baker* used the same hashtag for many of his tweets and it could be considered as his signature. *Tesla Motors* did not make much use of hashtags and their account relies on links and images instead.

### 8.2.1 Types of Tweets

The account of *Mark Baker* was divided in two almost equally large parts containing his own tweets and retweets, respectively. At first, I regarded it as a problem but in the end, it proved to be an advantage. This meant that I analysed tweets from more than five accounts and gained a greater insight. *The Telegraph* published only its own content and

did not share anything by other publishers. *Tesla Motors* retweeted posts from its CEO, Elon Musk, and from other accounts dedicated to energetics. Sharing related content by other authors was a common strategy for all accounts (except *The Telegraph*).

When a quote tweet is created, it automatically offers the option to add 140 more characters to the existing tweet. Retweets with an additional comment by the account owner were used by *Blake Harrison*, *Mark Baker* and *Oxford University*. I did not notice any endorsed content, i.e., retweeting sponsored advertisements.
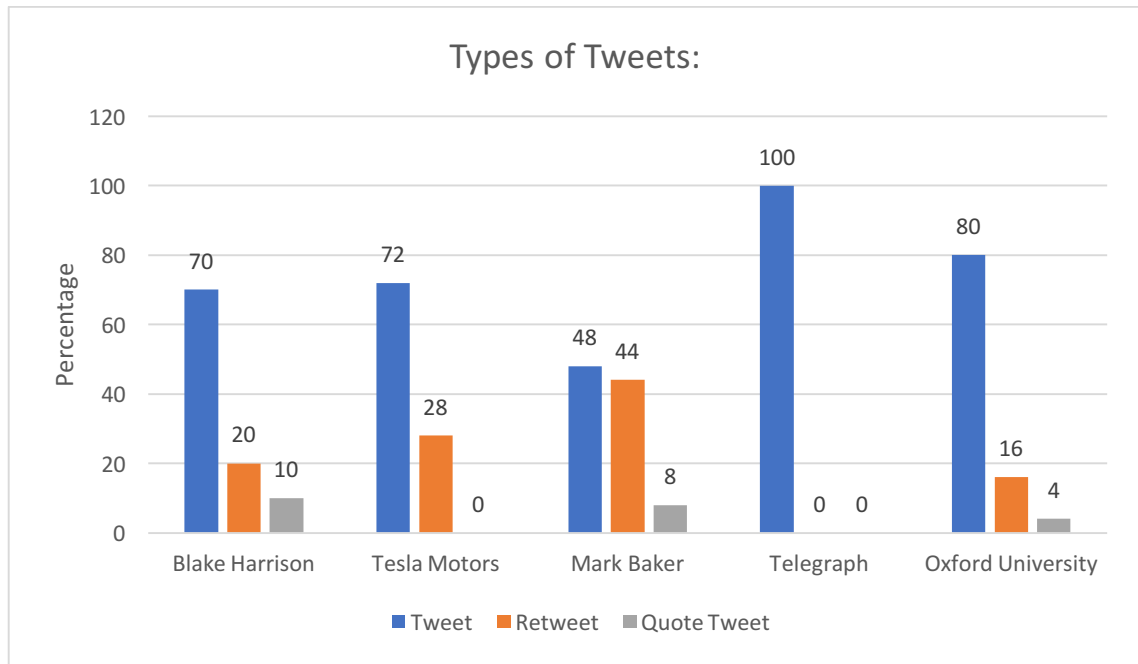


*Table 5: Types of Tweets*

## 8.2.2 Average Usage of Characters in Percentages

The graphics above shows the percentages of original tweets by the account owners versus retweets and quote tweets. *Mark Baker's* content and his retweets were long and used over 79% of maximum allowed characters. *Tesla Motors* did not need to reduce the number of characters as they used almost 57% of allowed space on average. The account who used the most space was *Oxford University*. It was due to the high number of links and hashtags.
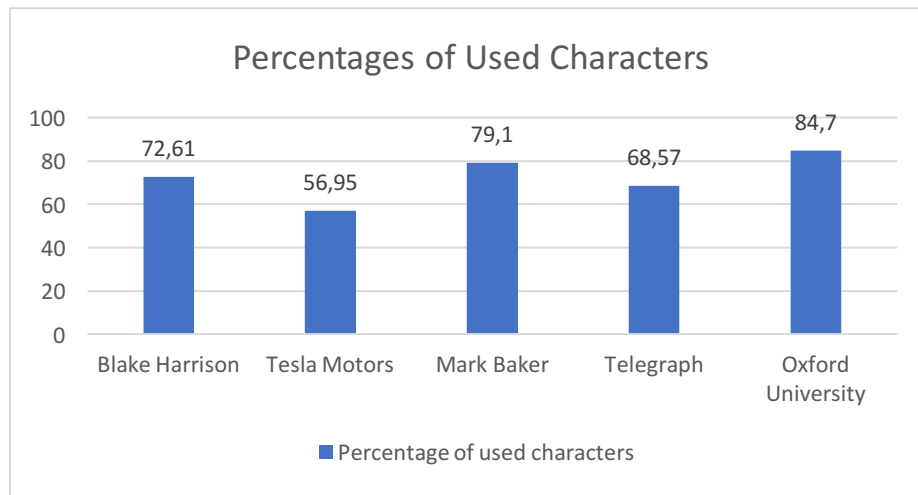
*Table 6: Percentages of Used Characters*

## 8.3 Sociolinguistic Aspects

All the analysed Twitter accounts are publicly accessible. This fact restricts their authors in terms of the content they can post. When posting inappropriate texts or images, authors may lose the trust of their followers. Blake Harrison is a 31-year-old man and shares content focusing mostly on his appearances on British series, TV shows and also private photos of his hobby – football. In this case, his tweets seemed appropriate to his age, occupation and social status. Mark Baker is a journalist and travel writer. This is why most of his tweets contain a photo and a comment on some place in a foreign country. He does his job and reports on it in his tweets. The content of other accounts is created by anonymous editorial staff and could not be analysed.

## 8.4 Conclusion

Despite its restriction in the maximum allowed characters and the lack of options for text formatting, Twitter offers interesting features like hashtags, retweets, media content and links. The behaviour of each analysed account was different. The institution account – Oxford University – referred to other sources of information through its tweets and added images to attract attention. Its posts were the longest ones.

Personal accounts were used for sharing their authors' own content and to promote their authors' work. Blake Harrison informed his followers about his appearances in other TV shows and Mark Baker shared his photos and knowledge of different parts of the world. He also retweeted many posts about travelling. Tesla Motors also retweeted posts from other accounts and more than half of their tweets contained a link. The Telegraph went straight to its own news and added images and hashtags to attract its followers.

### 8.4.1 Examples

Blake Harrison published tweets about his life, hobbies and appearances on British TV shows. Some of his posts:

"Air steward: would you like any crisps, popcorn or biscuits? Me: could I have all of them please? Air Steward: ...ok. Thug Life"

"My favourite #pancakeday recipe: Get a pancake Wrap it around a Twix Microwave for 20 seconds Eat with cutlery because you're classy."

Mark Baker added tweets with his own hashtag to make them easily identifiable across Twitter:

"On the way out of Astrakhan and back to Prague via Moscow. Wonderful trip through Russia for Lonely Planet. I'll be back in June #LPinRussia"

Oxford University dedicated its posts to topics from science and culture with links to articles:

"Happy #Dallowday! Read @OxUniPress's blog, celebrating Virginia Woolf: author, publisher, #feminist: http://po.st/ccrlol"

"We are talking about extending lives in a way we have never experienced before": Sarah Harper speaks to @guardian: http://po.st/kZK6aX

Tweets by Tesla Motors were simple (sometimes with an image or a video) and reported news about the company:

"Charging is our priority. This year we're doubling our Supercharger network & adding more chargers to existing sites [link]"

"Model S or Model 3: Which Tesla is right for you [link]"

# 9 Internet Forums

Internet forums are sometimes called bulletin boards, just boards or forums. They are sites with a community of users who are signed up and can express their thoughts about a variety of themes. Forums often split into several categories. The whole forum is usually structured into more specific groups of subjects for each topic which contain the individual posts. All posts must observe the rules of the whole forum and the authority who controls compliance with the rules is called an *administrator* or a *moderator*. There is more than one administrator responsible for the forum. Important topics from administrators are called *sticky topics* and are located at the top of the list of all topics to be visible for everyone and easy to find.

Before joining in the discussion, it is necessary to sign up for the forum. Membership is often free. After completing the registration process and logging in, the user can choose a topic to comment on or create his own topic. When no option to post a comment shows up (mostly a button called *Post Reply*), it may mean that the topic is closed for comments, the user was logged out or there is problem with permissions. The button *Quote* enables the user to make a more specific reaction to other posts. The post to which the user is responding is shown in a box and may be surrounded with quotation marks.
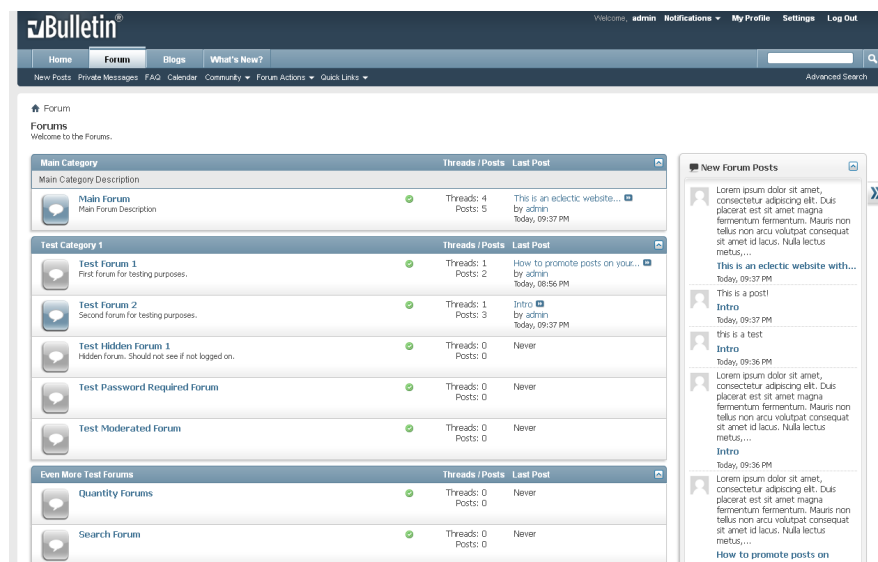


*Image 11: A Typical Design of an Internet Forum*

Users discuss under their nicknames, which they choose when signing up for the service, and are allowed to upload an image representing their account. This image is

called an *avatar*. Some forums do not allow to upload one's own image and the user is given the option to choose from an image gallery.

Any user can create his own topic. The topic must be created in the relevant category. To start with, the topic needs a suitable title. This title should attract the attention of other members. The space for writing posts is located under the title and there are often options to edit the text. To highlight some part of the content, the user can style the text in bold, change text size or colour, add emoticons, links, images etc. For making lists, there are bullets. After finishing writing and editing, there is a *Post* or *Send* button, which sends the post to the server of the forum and anyone can read and comment on the post.

Forums provide their users with the possibility to contact any other user via personal messages. This is achieved by clicking on a specific user name, after which his/her public profile page appears. There should be the button *PM* (Personal Message) or *Send a Message* (vBulletin).

## 9.1 Types of Forums

Since there is a huge number of possible topics to discuss, different types of forums are available. Some of them are dedicated to entertainment (games, movies, music etc.). It is also possible to find forums devoted to more specific themes (cars, fitness, real estate etc.). Besides these, there are discussion sites about any topic (Voat.co, Reddit.com and others). This type of forums usually lacks the classic structure of categories. Topics are posted on the main paged and sorted by votes from members.

People visit forums when they need to find information, sell or buy, express their opinion etc. In the modern world of the Internet, there are forums about almost any subject. Every forum has its rules and in the case of some specific forums (i.e., criminal activities, adult content etc.), these rules are more strict. There is also a huge number of abandoned forums on the Internet with no active users.

## 9.2 Highlighting Posts

Forums differ from one another also by different options for text formatting. The most frequently used variant is enclosing the text with predetermined symbols or code which change parameters like colour, font size etc. Some forums make use of *WYSIWYG*

editors.[8] When writing a text in a *WYSIWYG* editor, its formatting appears exactly as it will look like after posting. Emoticons are usually offered visually and can be added with just one click.

---

[8] "What You See Is What You Get" – an editor that allows the user to see what the end result will look like while the interface or document are being created.

# 10 Analysis of Internet Forum Posts

## 10.1 Introduction to the Analysis

I took a lot of time examining the different types of forums to find out how the topic categories are sorted and which topics to choose for analysis. In the end, I marked down 302 posts taken from 45 topics. A total of 101 posts came from GaiaOnline, 101 posts from Automotive Forums and 100 posts from Reddit.

## 10.2 Occurrences of Typical Features

My analysis revealed differences related to the length of the posts. Contrary to my expectations, more frequent occurrences of any feature were found in shorter posts. Shorter posts were more likely to contain any analysed attribute.
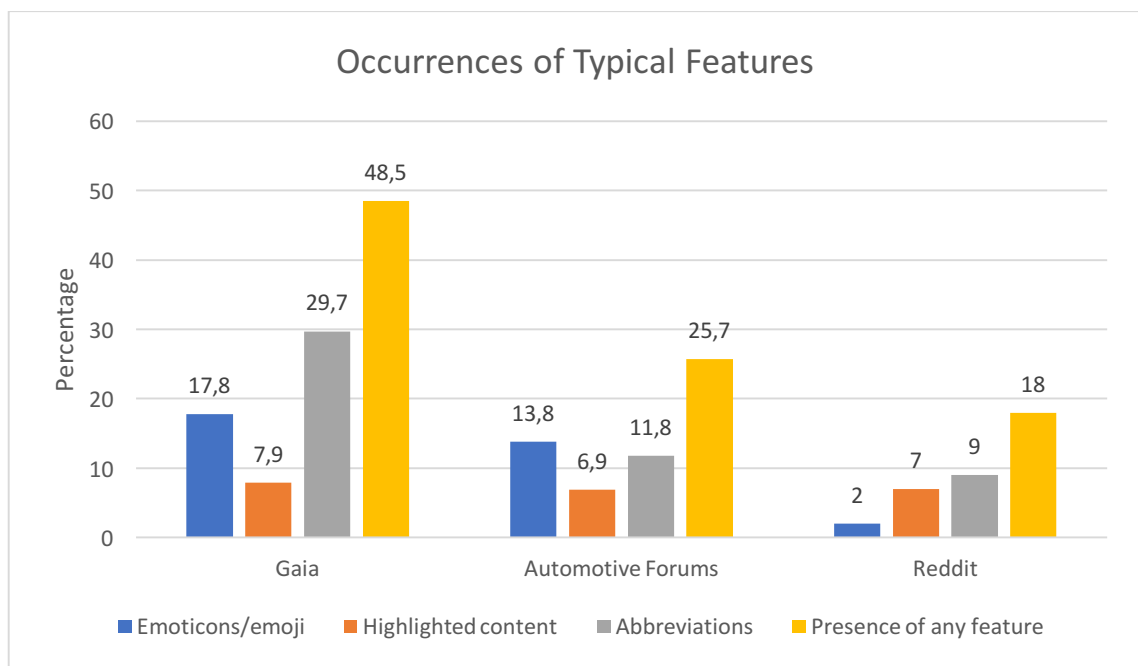


*Table 7: Occurrences of Typical Features of Forums*

Users of GaiaOnline were the most active group as to using emoticons and abbreviations and they also often highlighted their texts by changing the font colour. Frequent topics of GaiaOnline included a classic discussion of movies and books, fewer topics revolved around asking for and giving advice. On Automotive Forums, the members often asked for advice and offered advice to others. This led to less frequent occurrences of any feature. Members wrote longer posts to explain their answers more accurately, nevertheless, they still used emoticons and highlighted their posts. Reddit

manifested a comparable amount of highlighting but other features were present in low percentages.

### 10.2.1 Relevance of Comments

During my research, I realised that some comments were not relevant to the initial post. When the post was asking for advice, not every comment was useful, some of them were irrelevant and even off-topic. Reddit contained more than 60% of irrelevant comments. I sorted comments "by best" and analysed them starting from the top. As each comment reacted to the foregoing, they gradually lost their connection to the original post. While in classic forums, a post receives five to six replies on average, Reddit posts gained hundreds of replies.
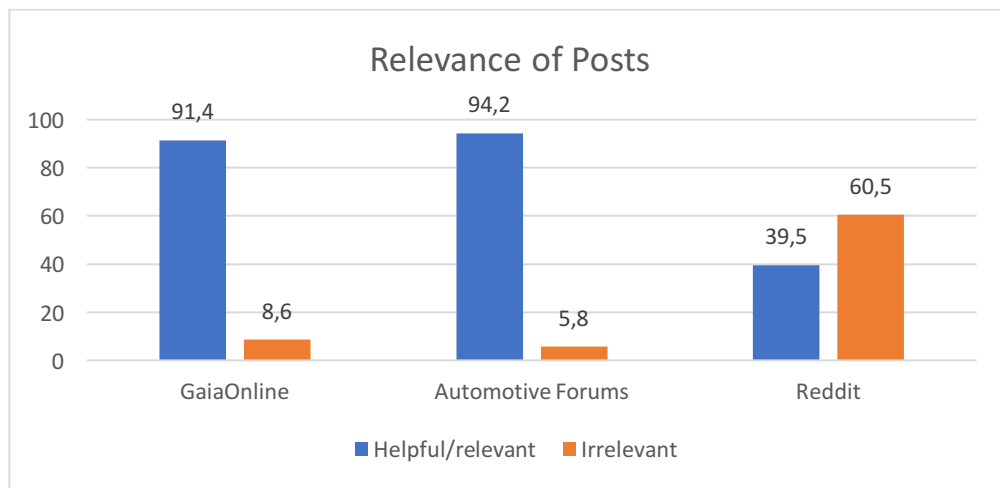


*Table 8: Relevance of Posts*

## 10.3 Sociolinguistic Aspects

It seems that Internet forums no longer cater primarily to children and young adults, who instead spend their time by using social networks. GaiaOnline targets at younger visitors by integrating mini games and topics relating to entertainment. The analysed posts however showed a relatively mature manner of expression. The users' nicknames were confusing and their avatars were taken from the gallery of the forum. I must admit that I expected more childish expressions and even more emoticons.

Young people who wanted to buy their first car often asked for advice on Automotive Forums. Their requirements were not very specific and they were apparently confused about more detailed advice from others. As to their avatars, they sometimes used a photo of their car.

Reddit was full of different opinions. Some content was marked as "18+" or as "NSFW" (Not Safe for Work), indicating that users might prefer not to visit this site during their working hours.

I did not notice any major grammatical errors or problems with expressing one's thoughts. The users communicated more or less sensibly, published their posts to the right category and formulated their texts comprehensibly.

## 10.3.1 Sarcasm and Politeness

Prior to starting the analysis, I had expected a high frequency of sarcasm and irony as well as impoliteness, but the results surprised me.



Quote:

Originally Posted by **Shaguar47**
*Dodge neon and Focus are good cars!*

Well im glad that the top all american sport cars made your list.
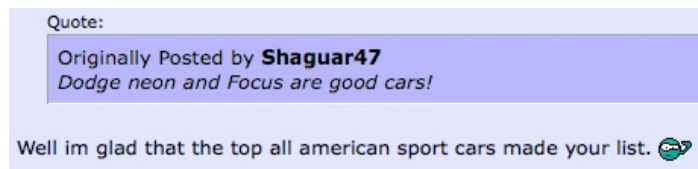
*Image 12: An Ironic Comment on Automotive Forums*

Only 3% of all analysed posts on GaiaOnline were impolite. The rest was mostly neutral. I had originally marked down occurrences of posts containing "thank you" and had classified them as polite, but then I decided to single out impolite comments only. Sarcastic remarks enjoyed a great visibility on Reddit. As the comments became less relevant, sarcasm grew more prominent.
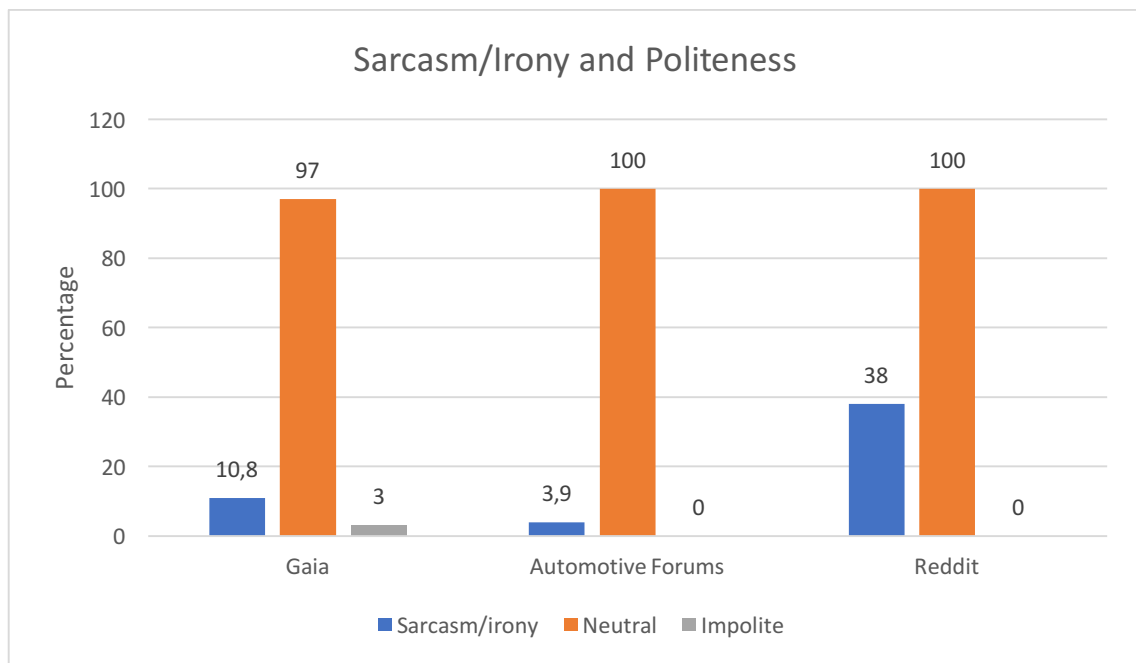


*Table 9: Sarcasm/Irony and Politeness of Posts*

45

## 10.4 Conclusion

Internet forums are characterised by a variety of users, each with their specific style of writing. Everyone can use the options for text formatting to attract attention to his post. Forums simplify the writing itself by allowing the users to work with editors, which supports creativity when preparing the text. The language on Internet forums is more complex because members of the forums discuss separated topics and need to be relevant. Reddit has a different way of a sorting topics and replies on them. This approach harms the consistency of the theme and results in a number of irrelevant comments.

The results surprised me with respect to the behaviour of the users. Sarcasm was still present, however, offensive comments with hateful messages were quite rare. GaiaOnline and Automotive Forums represented the classic forum system. Their posts contained more features typical for CMC.

### 10.4.1 Examples

The development of posts on Reddit shows the loss of connection to the original post:

- [the original post was a video of a pit stop in Formula One]

- IcePick1123: "I don't know if I'm more impressed with the pit crew, or the drivers ability to stop on a dime and not hit any of the pit crew"

- Lukebr4: "I know right! I'm sure the amount of time they spend In those cars he'll have better control over it than most people have over their own limbs."

- syndre: "they don't spend as much time in the car as one would think. They have a few weeks in the off season to test, a few hours of practice before each race…"

# 11 Comparing the Results

In this chapter, I would like to compare the obtained data and findings following from the results and evaluate the language typical for each of these modern Internet services.

## 11.1 The Focus of Each Service

YouTube comments are not so likely to get a response and their point is mostly expressing the thoughts of their authors. Some of the comments react to a particular moment in the video, some respond to the video as a whole. Many highlights and emoticons make each opinion stand out. The YouTube comment section is an integral part of YouTube itself but is not vital for the social network. Opinions are presented but they rarely lead to a more sustained conversation.

Twitter users post their tweets with the expectation to attract the interest of followers. Tweets often serve for sharing the original content of the author. This form of sharing is simple, easy to understand and authors typically write brief messages about their own activities. Sometimes, photos are attached. The accounts of institutions, brands and news publishers have a slightly different purpose. These accounts need to express their message so as to attract the reader's attention and offer a link to their website or another source.

Posts on Internet forums are typically discussions involving several participants. Whether it is a friendly discussion or a question for advice, it is expected that more users will join and add their own posts. The forum itself is the oldest service in my study which contributed to the creation of CMC to a large extent. Some forum topics may stay without any reply, others generate hundreds or even thousands of other posts and become a confused collection of opinions of many users.

## 11.2 Results in Numbers

These results are calculated from all the comments and posts that I collected (i.e., 950 YouTube video comments, 250 tweets and 302 posts from the Internet forums) and converted to per cent. The results may have been different if the numbers of texts had been the same for all services, but I believe that these results still well illustrate the basic idea.

- Emoticons were part of 11.5% of YouTube comments and 11.2% of posts from the Internet forums. Only 4.3% of all tweets contained an emoticon.

- Highlighted content appeared mostly in YouTube comments, 19.6% of them were somehow emphasised. The Internet forum posts used highlighting in 7.2%. On Twitter, only 2% of posts were characterised by highlighting.

- Abbreviations are typical for the forums with the proportion of 16.8%. Twitter is not far behind, as 9.2% of occurrences were found. YouTube comments did not contain these features that much, they amounted to 5.1% only.

## 11.3 The Main Purpose of Comments

On Twitter, authors of tweets comment on topical issues but are also allowed, for example, to show a picture of their new shoes. The authors wish to share their experience of the world and plan their tweets in advance. They are more responsible for their posts, as there are followers who expect interesting content and may leave when the content does not meet their expectations or standards.

YouTube users watch their favourite videos and may leave a comment to share their opinions or feelings. Users may however comment without seeing the video, which leads to offensive and impolite messages caused by a lack of interest, as they comment on something they hate. Comments on the largest video service on the Internet are characterised by many text highlights, which is the purpose of the users in most cases – to make their opinions visible.

Giving advice, asking for advice and discussing a common interest with others are the purposes of the forums. Like on Twitter, forum members plan their posts in advance when they are the authors of the first post or they react with their replies to an existing post. There is a risk of a poorly formulated question or a carelessly expressed opinion and some member may respond aggressively and add a hateful comment. Forum rules allow sarcasm but hateful content is punished.

# Conclusion

Judging from the results, it can be concluded that my research into the language of Internet forums, tweets and YouTube video comments has successfully fulfilled its purpose. I have analysed a total of 1 502 comments/posts and discovered the typical features which each service is characterised by. I have acquainted myself with many interesting passages in books and online sources concerning CMC, the Internet, its specific features and also sociolinguistics.

The language of the analysed Internet services varies, depending mostly on the purpose of the text. There are differences in whether a comment is expected to get a reply or whether it was created only to satisfy its author by sharing his opinion with the public. Sharing without expectations of developing a discussion is a matter of YouTube and Twitter. On Twitter, however, the author may expect some response from his followers.

 The language is also influenced by the presence of an audience. Each service has its typical features and the community uses them, but as the results of my study demonstrate, the frequency of their occurrence varies. The purposes of writing a post on an Internet forum, composing a tweet or leaving a YouTube video comment follow from different expectations of their authors. To generate discussion, people usually use the Internet forums, where they also ask for or give advice.

With respect to sociolinguistics, the final results remained within the norm of sociolinguistic groups. The reactions in selected texts were not so offensive and aggressive as I had expected, although sarcasm and irony appeared. It shows that users got used to enriching their expressions by adding any features of CMC. All comments and posts analysed are available for inspection on the CD enclosed with this thesis.

Several other questions remain to be resolved and I would like to continue in this topic in future and devote more attention to analysing further parameters of computer-mediated communication. There are still many opportunities to explore and many language phenomena of the web to compare. I am considering using this work as the basis for my future master's thesis.

# Works Cited

Abraham, Lee B., and Lawrence Williams. *Electronic Discourse in Language Learning and Language Teaching*. Philadelphia: John Benjamins, 2009. Print.

Alesso, H. P., and C. F. Smith. *Connections: Patterns of Discovery*. Hoboken, N.J.: Wiley-Interscience, 2008. Print.

Baron, Naomi. "Language of the Internet." *The Stanford Handbook for Language Engineers*. Ed. Ali Farghali. Stanford: CSLI Publications, 2003. 59–127. Print.

Barton, David, and Carmen Lee. *Language Online: Investigating Digital Texts and Practices*. Abingdon: Routledge, 2013. Print.

Boardman, Mark. *The Language of Websites*. London: Routledge, 2005. Print.

Boxer, Diana. *Applying Sociolinguistics*. Amsterdam: J. Benjamins, 2002. Print.

Brown, Steven, and Salvatore Attardo. *Understanding Language Structure, Interaction, and Variation: An Introduction to Applied Linguistics and Sociolinguistics for Nonspecialists*. 2nd ed. Ann Arbor: University of Michigan Press, 2005. Print.

Cantoni, Lorenzo, and Stefano Tardini. *Internet*. New York: Routledge, 2006. Print.

Crystal, David. *Internet Linguistics: A Student Guide*. Abingdon: Routledge, 2011. Print.

Crystal, David. *Language and the Internet*. 2nd ed. New York: Cambridge University Press, 2006. Print.

Danesi, Marcel. *Language, Society, and New Media: Sociolinguistics Today*. New York: Routledge, 2015. Print.

Danesi, Marcel. *The Semiotics of Emoji: The Rise of Visual Language in the Age of the Internet*. London: Bloomsbury, 2016. Print.

December, John. "Notes on Defining of Computer-Mediated Communication." *CMC Magazine* 4.1 (1997): www.december.com/cmc/mag/1997/jan/december.html. Web.

Dixon, Herbert. "Texting, Tweeting, and Other Internet Abbreviations". *The Judges' Journal* 50.4 (Fall 2011): 30–33. www.americanbar.org/content/dam/aba/publications/judges_journal/2011falljj_t ech.authcheckdam.pdf. Web.

Holmes, Janet. *An Introduction to Sociolinguistics*. New York: Longman, 1992. Print.

Holt, Richard. *Dialogue on the Internet: Language, Civic Identity, and Computer-Mediated Communication*. Westport, Conn.: Praeger, 2004. Print.

Livnat, Zohar. "On Verbal Irony, Meta-Linguistic Knowledge and Echoic Interpretation." *Pragmatics & Cognition* 12.1 (2004): 57–70. Print.

Meyerhoff, Miriam. *Introducing Sociolinguistics*. London: Routledge, 2006. Print.

Posteguillo, Santiago. *Netlinguistics: An Analytical Framework to Study Language, Discourse and Ideology in Internet*. 2. ed. Castelló de la Plana: Publicacions de la Universitat Jaume I, Servei de Comunicació i Publicacions, 2003. Print.

Sahlin, Doug, and Chris Botello. *YouTube for Dummies*. [Chichester: John Wiley], 2007. Print.

Salvatore, Attardo. "Irony as Relevant Inappropriateness." *Journal of Pragmatics* 32 (2000): 793–826. http://omero.humnet.unipi.it/matdid/372/Attardo_Irony.pdf. Web.

Shortis, Tim. *The Language of ICT: Information and Communication Technology*. London: Routledge, 2001. Print.

## Internet Sources

Adib, Desiree. "Pop Star Justin Bieber Is on the Brink of Superstardom." *ABC NEWS*. ABC News Internet Ventures, 14 Nov. 2009. www.abcnews.go.com/GMA/Weekend/teen-pop-star-justin-bieber-discovered-youtube/story?id=9068403. Web. June 20, 2017.

Alexa. "Reddit.com Traffic Statistics." Alexa Internet, n.d. www.alexa.com/siteinfo/reddit.com. Web. May 28, 2017.

AutomotiveForums. "Corporate." Automotive Network, n.d. www.automotiveforums.com/corporate/index.php. Web. May 28, 2017.

Business Dictionary. "Internet – Definition." WebFinance, n.d. www.businessdictionary.com/definition/internet.html. Web. May 28, 2017.

Crystal, David. "How Is the Internet Changing Language Today?" *YouTube*. Macmillan Education, 20 Jan. 2010. www.youtube.com/watch?v=P2XVdDSJHqY. Web. May 28, 2017.

Emojipedia. "FAQ." Emojipedia, n.p. www.emojipedia.org/faq/. Web. Apr. 6, 2017.

English Oxford Living Dictionaries. "Abbreviation." Oxford University Press, n.d. www.en.oxforddictionaries.com/. Web. May 26, 2017.

English Oxford Living Dictionaries. "Acronym." Oxford University Press, n.d. www.en.oxforddictionaries.com/. Web. May 26, 2017.

English Oxford Living Dictionaries. "YouTuber." Oxford University Press, n.d. www.en.oxforddictionaries.com/. Web. May 26, 2017.

F Newsroom. "Our Mission." Facebook, n.d. www.newsroom.fb.com/company-info/. Web. Feb. 1, 2017.

GaiaOnline. "About Gaia, Interactive, Inc." Gaia Interactive, n.d. www.gaiaonline.com/info/about/. Web. May 28, 2017.

Google Support. "Age Restrictions." *Google Account Help*. Google, n.d. www.support.google.com/accounts/answer/1350409?hl=en. Web. May 28, 2017.

Internet World Stats. "Internet Usage Statistics." Miniwatts Marketing Group, n.d. www.internetworldstats.com/stats.htm. Web. Feb. 6, 2017.

Jack. Image Tweet. *Twitter*. Twitter, Jan. 5, 2016.www.twitter.com/jack/status/684496529621557248. Web. May 1, 2017.

Jokinen, Jari. *Percentage Calculator*. Jari Jokinen, n.d. www.percentagecalculator.net. Web. May 28, 2017.

Ludwig. *Ludwig*. Ludwig, n.d. www.ludwig.guru. Web. May 28, 2017.

McCoy, Simon. "Simon McCoy Blooper – Drops His Career on the News!" *YouTube*. MonaichFother, Jan. 21, 2016. www.youtube.com/watch?v=K41pTIrsNgg. Web. June 23, 2017.

Moritz, Donna. "Twitter Gets Visual – 3 Image Mistakes You Are Making on Twitter." *Socially Sorted*. Socially Sorted, 2014. www.sociallysorted.com.au/twitter-gets-visual-3-image-mistakes-you-are-making-on-twitter/. Web. May 2, 2017.

Stands4. *Abbreviations*. Stands4, n.d. www.abbreviations.com. Web. May 28, 2017.

Titcomb, James. "Twitter Now Lets You Post Longer Tweets after 140-Character Limit Relaxed." *The Telegraph*. Telegraph Media Group, Sep. 20, 2016. www.telegraph.co.uk/technology/2016/09/20/twitter-now-lets-you-post-longer-tweets-after-140-character-limi/. Web. March 2, 2017.

Turner, Amy-Mae. "9 Breaking News Tweets That Changed Twitter Forever." *Mashable*. Mashable, Oct. 31, 2013. www.mashable.com/2013/10/31/twitter-news/#Vt1Wb._M0aqE. Web. March 2, 2017.

Twitter. "Twitter Milestones." *About Twitter*. Twitter, n.d. https://about.twitter.com/company/press/milestones. Web. Feb. 1, 2017.

vBulletin. "Forums, Topics and Posts." vBulletin, n.d. www.vbulletin.com/forum/help?faq=vb3_board_usage#faq_vb3_forums_threads_posts. Web. June 5, 2017.

Wall, Aron. "Why Are Internet Discussions Less Polite?" *Undivided Looking*. Aron Wall, May 21, 2015. www.wall.org/~aron/blog/why-are-internet-discussions-less-polite/. Web. May 28, 2017.

Wikipedia. "Reddit." Wikimedia Foundation, n.d. www.en.wikipedia.org/wiki/Reddit Web. May 28, 2017.

Yoree, Koh. "What Twitter Gains by Expanding the 140-Character Tweet." *The Wall Street Journal*. Dow Jones Company, June 1, 2016. https://blogs.wsj.com/digits/2016/01/06/what-twitter-gains-by-expanding-the-140-character-tweet/. Web. Apr. 2, 2017.

YouTube Help. "Post & Interact with Comments." *Support Google*. Google, n.d. www.support.google.com/youtube/answer/6000964?hl=en&ref_topic=3014329. Web. May 28, 2017.

YouTube Help. "View, Organize, or Delete Comments." *Support Google*. Google, n.d. www.support.google.com/youtube/answer/6000976?hl=en. Web. May 28, 2017.

## Images

"The Main Page of Twitter." Twitter, 2017. Author's screenshot. www.twitter.com.
    Web. May 23, 2017.

 "The Main Page of YouTube with Popular Videos." YouTube, 2017. Author's
    screenshot. www.youtube.com. Web. June 26, 2017.

"vBulletin." vBulletin, 2017. Author's screenshot.
    www.files.vbulletin.com/images/cart/screenshots/forum_main.png. Web. June
    14, 2017.

## List of Images

## List of Tables