LINGUISTIC VARIATION ACROSS TWITTER AND TWITTER TROLLING

by

ISOBELLE CLARKE

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Department of English Language and Applied Linguistics
University of Birmingham
January 14th, 2020

Abstract

Trolling is used to label a variety of behaviours, from the spread of misinformation and hyperbole to targeted abuse and malicious attacks. Despite this, little is known about how trolling varies linguistically and what its major linguistic repertoires and communicative functions are in comparison to general social media posts. Consequently, this dissertation collects two corpora of tweets – a general English Twitter corpus and a Twitter trolling corpus using other Twitter users' accusations – and introduces and applies a new short-text version of Multi-Dimensional Analysis to each corpus, which is designed to identify aggregated dimensions of linguistic variation across them. The analysis finds that trolling tweets and general tweets only differ on the final dimension of linguistic variation, but share the following linguistic repertoires: "Informational versus Interactive", "Personal versus Other Description", and "Promotional versus Oppositional". Moreover, the analysis compares trolling tweets to general Twitter's dimensions and finds that trolling tweets and general tweets are remarkably more similar than they are different in their distribution along *all* dimensions. These findings counter various theories on trolling and problematise the notion that trolling can be detected automatically using grammatical variation. Overall, this dissertation provides empirical evidence on how trolling and general tweets vary linguistically.

Acknowledgements

This dissertation would not have been possible without the support and guidance of my supervisor Professor Jack Grieve. Jack has given so much time to me. He has provided me with food for thought and helped me to learn programming and always managed to see the bigger picture beyond the thesis, so thank you Jack. You have been an incredible supervisor and you've got me set for a prosperous career... Well, however prosperous academia can get me.

This dissertation began at the Centre for Forensic Linguistics at Aston University and so I would like to thank everyone there, especially Tim Grant, Krzysztof Kredens, Annie Heini, Emily Carmody and Dan Thompson for all of your support.

My sincere thanks goes to everyone in the Department of English Language and Applied Linguistics at the University of Birmingham who listened to my presentation and offered suggestions and feedback, which have certainly informed this dissertation. Special thanks goes to Bodo Winter for giving me the experience to teach MCA to others - my maths teacher at school would be so proud!

I would also like to thank Aston University and the college of Arts and Law who funded this thesis and my life for the past 3 years, enabling me to dedicate all my time to my thesis... Some people are not afforded the same opportunity and I am truly grateful.

I would also like to thank my family and friends for always believing in me when I didn't myself, and offering support when I needed it. I've not always been the most pleasant, fun, attentive or caring throughout this process, so thank you for being so understanding.

Finally, I would like to thank my partner Mark McGlashan for all of his love, support and patience. Thank you for allowing me to talk *at you*. Thank you for understanding when I'm being difficult and thank you for giving me advice even when I said I didn't need it. It was *always* appreciated. I'm not sure anyone else would have put up with me... I've certainly lost my patience with myself. Thank you for allowing me to do *me*, and at the same time being the best kind of distraction. I really do love you.

Despite a thesis in Linguistics, there really aren't the right words to express how appreciative I am of all of you along the way, so let's just imagine ‘thank you’ suffices. Thank you.

Contents

List of Figures

List of Tables

List of Abbreviations

CMC – Computer-Mediated Communication

CoNLL - Computational Natural Language Learning

MCA – Multiple Correspondence Analysis

MDA – Multi-Dimensional Analysis

MP – Member of Parliament

NLP – Natural Language Processing

URL – Universal Resource Locator

1. Introduction

When I was at school, there was a young boy who was exceptionally clever – he always had the answer or a smart remark to offer, and continuously corrected other students and also the teachers. But the students did not appreciate this ability and poked fun at him and he was often the target of teasing and abuse. One day, when the teasing became too much, he erupted – he threw chairs, swung his bag around and attempted to fight back at the bullies, raising his fists and swinging. His outrage did not stave off further teasing. Instead, his loss of control became a laughing matter, a punch line, and suddenly, a game. The reaction - the loss of control, and the unruly outbursts of frustration was something that his classmates and students in other classes sought to provoke. It was relentless. They said things they knew would get a rise out of him, even though they did not mean it, and they did it all for entertainment – because it was funny.

Getting a rise out of someone or provoking a negative reaction for entertainment permeates the offline world from the school playground to political debates and elections. And with the creation of the Internet and the World Wide Web and the rise of social media, it has spread to the online spaces and developed a new name - *trolling*. Trolling involves behaviours that are purposefully designed towards triggering a predominantly negative reaction from a particular individual, group or community. Because being provoked is a personal and subjective thing - what provokes you may not necessarily provoke me - trolling has developed into a multi-faceted phenomenon. Although research acknowledges its different guises and has begun to detail some of its behaviours and characteristics, little is known about the communicative functions of trolling and how it varies linguistically. For instance, whilst trolling can involve various behaviours, such as malicious attacks, false information, false advice, hostile content, exaggerations, mocking and teasing, there lacks an understanding about the major linguistic repertoires and properties of trolling. Importantly, the question arises about whether such behavioural distinctions of trolling are reflected in real linguistic differences across trolling instances. This dissertation begins to fill that gap by conducting the first linguistic investigation into the major communicative functions of trolling as carried out on Twitter.

Twitter is a social media platform, which grew out of the shift from web 1.0 to web 2.0, and the bursting of the dotcom bubble (described below). Whilst many of the social media platforms created during this time, including Twitter, were aimed at bringing people closer together, they have nevertheless become spaces for trolls to invade and disrupt. The following section tracks the development and rise of social media platforms, especially Twitter, and its various technological features and affordances.

1.1. The Internet and the Rise of Social Media

The Internet was originally designed to transfer information and communicate with other computers in the scientific domain for intelligence and defence purposes (see Herring, 2002). At this time, computers were limited to specialist use due to their cost and the need for technical expertise (Thurlow, Lengel and Tomic, 2004). Then, with the invention of the World Wide Web by

Tim Berners-Lee, the Internet became a place where information and content was stored and could be accessed and consumed far easier via websites and hyperlinks (Seargeant and Tagg, 2014). In its beginning in the early 1990s, the World Wide Web was comprised of mainly static documents, lists of frequently-asked questions, e-commerce sites and personal homepages that were entirely connected by hyperlinks, and contained content that was for the purpose of reading (i.e. consumption) (Herring, 2012). For example, by the end of 1994, many websites that are popular nowadays already existed, including web portals like Yahoo!, online newspapers (e.g. Telegraph, The Irish Times), local government websites (e.g. Birmingham City Council), food delivery websites like Pizza Hut, fan sites (e.g. The Simpsons Archive), and sites for movies and music information (e.g. MTV, IMDB), among others. Although some e-commerce sites already existed, from 1995, the potential of commerce on the web was gaining more attention with many speculating that Internet based companies were going to be highly profitable (Hayes, 2019). Consequently, from 1995 to 2000, venture capitalists abandoned the cautious approach and instead poured free money into Internet start-ups with a ".com" site with the hope that they would return a profit (Hayes, 2019). This period is widely recognised as the dot-com bubble, where technology stock equity valuation rose rapidly, and it was fuelled by fad-based investing and speculation (Hayes, 2019). The dotcom bubble eventually burst in 2001 and through to 2002, where a huge number of internet-based companies went bust and investors faced considerable losses (Hayes, 2019). Amazon, eBay and Priceline are some of the companies that survived the bubble (Hayes, 2019).

Before the bursting of the dot-com bubble, during the late 1990s, there was also a shift from static websites to considerably more dynamic and interactive websites (Herring, 2012), where commenting and conversation were encouraged and fostered, resulting in content that was no longer consumed, but created by users. During this time, computers became easier to use and more affordable, and by extension, they became considerably popular, especially for the purpose of human interpersonal interaction (Thurlow, Lengel and Tomic, 2004), facilitated by the shift in increasingly more interactive websites. Although highly debated, this shift has been labelled as a move from Web 1.0 to Web 2.0 (O'Reilly, 2005), where the web became conceptualised as a product of participation as opposed to publishing. Some of the features of Web 2.0 sites include those that harness user contributions, as well as those that have software that can be used across more than one device (O'Reilly, 2005). For example, Amazon harnesses user activity, such as sales information, to display the most popular search results (O'Reilly, 2005).

One of the most distinguishing features of Web 2.0 sites is their focus on owning their own data set - one which is hard to recreate and unique - that gets richer as more people use them (O'Reilly, 2005), because data is a valuable resource, which can be bought, sold and used for profit. For example, the social network site, Facebook manages online personas (based on personal information, friends, likes, clicks on advertisements, etc.) and sells that information produced by users to marketing companies so that advertisements can be tailored to the right Internet user (Karppi, 2013). Similarly, it is reported that Twitter earned $2.61 billion in revenue from advertising, which is 86% of its gross revenue (Beers, 2019). Social media platforms are purposely designed for user participation, content creation and social interaction (Seargeant and Tagg, 2014). Thus, in this move to web 2.0, there was a clear increase in the creation and use of social media sites, which led to a rise in online interactivity and user participation (Seargeant and Tagg, 2014; see boyd and Ellison, 2007), and ultimately, the creation of mass datasets of user-generated content. Human participation online has become a commodity, where each individual leaves digital traces, which are being mined and amalgamated into an extremely high profit industry.

Social media is defined here broadly as any digital environment that involves human interaction (Leppänen et al., 2014; Seargeant and Tagg, 2014), allowing for user-generated content to be created and exchanged (Kaplan and Haenlein, 2010). Social media is therefore understood in this dissertation as facilitating and playing host to communication and interaction mediated through digital technology with a move from the traditional notion of 'content' being something that is published, broadcast and consumed (such as with Web 1.0 sites and traditional print and news media) to the idea that 'content' on social media is a product of participation, generated, developed and shared by and amongst users (see Seargeant and Tagg, 2014). SixDegrees.com was arguably the first social media site created in 1997, which encouraged users to create their own profile and add others to their personal network (Terrell, 2019). It was based on the six degrees of separation, where any individual in the world is connected to everyone else by no more than six levels of separation. As a result, the site connected individuals and revealed a network of relations to other individuals (Terrell, 2019). This site also enabled users to send messages to people in their first, second and third degrees of separation. SixDegrees.com only lasted until around 2001 (Terrell, 2019); nevertheless, new social media platforms were on the horizon. For example, Friendster emerged in 2002 (boyd and Ellison, 2007; Terrell, 2019). It began as a social-networking site, where users could make contacts and save them as part of a personal network (Terrell, 2019). Users could post comments to the profiles of the people within their network, as well as share messages, videos, and photos with other users (Terrell, 2019). Despite being rebranded in 2011 as a social gaming site as a result of fierce competition from other social networking sites, especially Facebook, it suspended all services on 1st January 2019 (Terrell, 2019). Other popular social media platforms that were founded at a similar time as Friendster include LinkedIn in 2002, MySpace in 2003, Facebook in 2004, Reddit in 2005, and Twitter in 2006 (Terrell, 2019). Instagram and Snapchat were launched later in 2010 and 2011, respectively (Terrell, 2019). Most social media platforms generally, can be used to keep up-to-date with friends, family, colleagues and other acquaintances, as well as complete strangers, celebrities and people of interest.

There are different types of social media platforms each with their own design and set of resources, including platforms designed for sharing photos (e.g. Instagram, Snapchat) and videos (e.g. YouTube, Vimeo), as well as platforms designed for social networking (e.g. LinkedIn) and for articulating and making visible one’s social network (e.g. Facebook) (boyd and Ellison, 2007). Additionally, there are platforms tailored for blogging and microblogging (e.g. Twitter, Tumblr), as well as collaborative writing projects, such as Wikipedia, discussion forums, such as Reddit and StackExchange, virtual game worlds (e.g. World of Warcraft, Second Life), and platforms dedicated to business and product reviews (e.g. Zagat and Foursquare) (boyd and Ellison, 2007). Notably, social media platforms can incorporate and display more than one of these design features. For example, Lee (2011) indicates that microblogging, which generally refers to short messages written on the web that self-report on what one is doing, thinking or feeling at a particular moment, can be performed on microblogging specific platforms, such as Twitter, as well as on social network sites like Facebook. In the same vein, Facebook and Twitter can also be used to share photos and videos.

Twitter, originally called ‘twttr’, was envisioned by Jack Dorsey as a short message service (SMS) based communications platform, where friends could keep up-to-date with what each other was doing based on their status updates called tweets (MacArthur, 2019). This idea was pitched to co-founders Evan Williams and Biz Stone, and together with Noah Glass, ‘twttr’ came into existence in 2006. The first tweet sent by Dorsey said: “just setting up my twttr” (@jack, 2006) (MacArthur, 2019) and the first tweet sent by one of the co-founders Biz Stone said: “Ok we are in the car” (@biz, 2006) (Zappavigna, 2018).

The completed version of Twitter debuted in March 2007 at the conference South by Southwest, where over 60,000 tweets were posted each day at the conference (MacArthur, 2019). Twitter grew quickly and by the first quarter of 2010, it had 30 million monthly active users (Clement, 2019). This number grew consistently up to the first quarter of 2017, where Twitter had 327 million monthly active users. Since then, the number of monthly active users has fallen and risen, and in the first quarter of 2019, Twitter had 330 million monthly active users from across the globe (Clement, 2019).

Users of Twitter must create a profile and can then begin to post tweets, which appear on one’s profile. Users can follow the tweets of others. Twitter is based on the practice of following, which is similar to subscribing to someone’s updates. Following is non-reciprocal on Twitter, which means that users do not have to follow back the people who follow them. The non-reciprocity of following leads to complex follower networks being formed, which include unidirectional and bidirectional connections with a variety of individuals, organisations (governmental and non-governmental), and media outlets (Weller et al., 2014). The tweets of the people that one follows appear in one’s timeline, which is a filtered version of the public stream of tweets according to who the user follows. Tweets can concern a whole range of topics, from politics to pizza, and a whole range of text types, from jokes to recipes, for a variety of different purposes, including entertainment and activism. The major restriction on tweets is on its size, which was initially constrained to 140 characters to accommodate the length of SMS texts, as this was the standard way to send tweets. However, in November 2017, tweet length was extended to 280 characters per tweet. Tweets can now be posted through the web, applications and third-party clients on phones, computers and other devices.

Given these design features of Twitter and its popularity, Twitter can be highly instrumental for a number of important tasks. Specifically, Twitter has been praised in its capacity to bring people from across the globe and from all walks of life closer together. For instance, Twitter provides an access channel to celebrities and they use the platform to communicate with fans. Additionally, Twitter gives people, especially silenced individuals and marginalised groups a voice, providing them with the opportunity to be heard. For example, Cui Haoxin, a Chinese Muslim poet spoke out on Twitter about the amount of Islamaphobia in China (Shih, 2019). Twitter also provides the tools to organise and orchestrate protests and political activism, and share important information across the world, such as with the "Arab Spring" uprisings in 2011 (Bruns et al., 2013) and disasters like the Pakistani floods in 2010 (Murthy and Longwell, 2013). Whilst these examples illustrate the advantages of Twitter and the positives of Twitter's design, there are some individuals who exploit these freedoms and provisions for negatively marked means, especially for the purpose of provoking a response. These individuals have been labelled in a variety of ways, but most commonly they are called *trolls and* their behaviour is deemed *trolling*.

1.2. What Is Trolling?

The terms *troll* and *trolling* mean different things for different people. Its first usage online can be traced back to Usenet in 1992, where it referred to the act of posting an exaggerated message on a previously discussed topic into the newsgroup (NetLingo, 1995-2015). The aim of doing so was to expose new members, as they would not be aware that it had been mentioned before, and would subsequently be provoked to post corrections or point out the misconception, ultimately revealing their 'newbie' status. Experienced members of the newsgroup would know that the original poster was not being serious and was essentially trolling and would consequently not respond. For example, one example of trolling involved an advert that instructed individuals to delete System32 off their computer in order to speed it up (Phillips and Milner, 2017). Doing so, however, turns the computer into a brick. New members would respond to warn other users of this misconception, whereas experienced members would not bite the bait and would allow the fake advice to persist.

There are two popular theories for the use of the terms *troll* and *trolling* on the Internet to describe this behaviour. The noun *troll* was originally used to describe ugly creatures in Scandinavian mythology and folklore (Harper, 2017). While there are several different depictions of trolls in these tales, they are predominantly framed negatively and can be tricksters who disguise their true self and appearance to manipulate humans into doing something (Ljosland, 2013). Alternatively, the verb *to troll* was initially used to describe fishing with a moving bait, which led to the figurative use of the verb to mean luring or enticing with a bait (Harper, 2017). Both of these uses of *troll* have an element of trickery through disguise, which connects to its application online to post some form of inaccurate or hyperbolic message, which is disguised as a genuine post to trick and provoke others into replying (or biting the bait).

Despite the original usage of *trolling* on Usenet – to provoke a response from new members in order to expose them – the term has spread across many platforms and sites, and has become associated with a much more general aim – to provoke and manipulate others into doing something, mainly to respond in some way. Given this aim, deception is largely a necessary part of the practice. Trolls must disguise their trolling intentions because if their victim becomes aware of them, then they will not engage, starving trolls of the reaction they crucially crave. Donath (1999) and Hardaker (2010) describe how important the process of deception is for trolling, as it increases the chance of achieving a reaction from their victim(s) and ultimately being successful. To disguise their intentions and provoke a response, the troll may have to observe the particular online community, so that they can behave similarly to appear like a genuine member (Donath, 1999; Hardaker 2010). Cruz et al. (2018) describe this as a process of learning and assimilation. Having learned and begun to assimilate towards the community, trolls can then transgress, posting comments and behaving in ways that are at variance with the community's standards and norms in order to provoke the response (Cruz et al., 2018; Phillips, 2016). Because each online community has their own norms and standards, trolling has become an umbrella term that encapsulates so many online behaviours that are transgressive for the particular community. Moreover, with the influence of the media, and their focus on exceptionally negative instances of online abuse, trolling has expanded to include behaviours that are prosecutable under the UK's communication acts.

The Communications Act (2003: section 127) makes it an offence to send messages or other matter via a public electronic communications network, which "is grossly offensive or of an indecent, obscene or menacing character". This particular act has been used to convict various individuals, depicted as 'trolls' in the media. For example, Isabella Sorley and John Nimmo were both sent to prison for sending menacing tweets to Caroline Criado-Perez, who was campaigning to put a female on the back of a bank note (Cockerell, 2014). Peter Nunn also received a prison sentence following his threatening messages to

Stella Creasy, a Labour MP, who was in support of Criado-Perez's campaign (BBC News, 2014). More recently, John Nimmo was sentenced again for his grossly offensive, threatening and purposely-false messages that he sent to Luciana Berger (a Labour MP) (Laville, 2017).

Although prosecuted cases of 'trolling' and those reported in the media tend to reflect the worst of the worst, research shows that not all cases of trolling are perceived to be negative (Sanfilippo, Yang and Fichman, 2017). In fact, some cases of trolling are perceived to be playful and ingenious. For instance, one common type of playful trolling is called Rick Rolling, which involves the troll posting a message, which subtly persuades you to click on a URL. This URL ends up being irrelevant to the message, and instead takes you to Rick Astley's music video "Never Gonna Give You Up". For example, there was once a Reddit post, where numerous Reddit users were discussing what trolling was. One poster provided their insight on trolling and its history. and subsequently supported their definition by describing how someone had given a TED talk on the subject and that this could be accessed here, with the word 'here' hyperlinked. The link turned out to be Rick Astley's "Never Gonna Give You Up" music video, as opposed to a TED talk on trolling. Hundreds of individuals all below the comment posted applause emojis, congratulating the poster for their seamlessness in trolling them all on, ironically, a thread about trolling. Such playful cases of trolling are not necessarily reported in the media. Instead, the media tend to describe more criminal and negatively marked behaviours as trolling. For example, media reports described groups of "racist trolls" that had targeted Leslie Jones on Twitter following the release of the new Ghostbusters film (Woolf, 2016). As a result of its diverse uses across online communities and platforms and in the media, trolling has developed into a term that, depending on the context, can mean anything from hate speech and malicious cyber-bullying to playful banter and pranks.

Using *trolling* to describe hate speech and other prosecutable behaviours is problematic, as it could lead to the desensitisation of hate speech. Specifically, one of the most common reported goals of trolling is to provoke a response, and consequently, there is the view that trolls will say and do anything to achieve this, even if they do not agree or believe what is being said. Isabella Sorley, for example, threatened Criado-Perez, indicating that Sorley would do a lot worse things than rape Criado-Perez. It is not clear what these ‘worse things’ were or whether Sorley would actually have done them – the point was to antagonise her and provoke a response, regardless of whether she meant it or not. Emphasising the end goal (i.e. provoking a response), however, means that the means by which the goal is achieved is arguably excused (i.e. “I’m not really racist, I just said racist things to provoke you”). For instance, one Reddit user called ‘HanAssholeSolo’, who had previously posted numerous bigoted, racist and anti-Semitic messages, posted a video, which was shared by Donald Trump (Gabbatt, 2017). When journalists sought to uncover ‘HanAssholeSolo’s’ true identity (which they did), it provoked ‘HanAssholeSolo’ to apologise not only for the crude video, which was shared by Trump, but also for the numerous bigoted messages. He said:

> *I would also like to apologize for the posts made that were racist, bigoted, and anti-semitic… I am in no way this kind of person, I love and accept people of all walks of life and have done so for my entire life. I am not*

> *the person that the media portrays me to be in real life, I was trolling and posting things to get a reaction from the subs on reddit and never meant any of the hateful things I said in those posts.*
>
> (HanAssholeSolo, 2017 cited in Gabbatt, 2017)

Many trolls, especially when faced with serious consequences like exposing their identity or prosecution, resort to defending their actions by insisting that they were trolling, and therefore they were only joking and did not mean it, rather they just meant to get a reaction. Whilst this offers a way to distinguish trolls from other kinds of bigots (i.e. true racists would not necessarily defend their statements by insisting they were just joking around), it is nevertheless problematic because it has been argued that the excuse of “I was just trolling” when articulating hate speech could lead to the desensitisation of hate speech (Phillips and Milner, 2017), and also that this excuse may even be used by racists as a kind of ‘get out of jail free card’ (Clarke, 2019). Whilst this is a potential consequence of trolling, the list of actual disastrous effects is growing.

In particular, reports exist of trolls sowing discord in the democratic process, especially by spreading fake news, misinformation and disinformation (Lewandowsky et al., 2017; de Quetteville, 2018). For example, an office block in St. Petersburg was exposed as a Russian Troll Factory, which was designed to pollute Twitter and other social media platforms with an abundance of disinformation in order to fray the fabric of Western Society throughout the 2016 US presidential elections (de Quetteville, 2018). The spread of disinformation is not necessarily aimed at persuading people, although this can happen, but rather it is aimed at complicating the context and knowledge by polluting it with an abundance of competing, false and useless information (Fokin, 2016). This can ultimately instil doubt in common knowledge and drive conspiracy. For example, there has been an ongoing disinformation campaign on vaccinations, where the positive, the negative and pseudo effects of vaccinations are articulated and spread (de Quetteville, 2018). Mahase (2019) has emphasised a global 300% rise in measles in the first few months of 2019, although the increase is higher and lower in some countries. In particular, Mahase (2019) reports that from 1 January 2019 to 11 April the US has seen the second highest number of annual measles cases since it was eliminated in 2000. Whilst trolling and the spread of disinformation is not the sole reason for the increase in measles cases, there is no doubt that trolling and disinformation campaigns have had an influence on this rising number (de Quetteville, 2018).

Moreover, there are reports of other disastrous consequences of trolling. For example, Charlotte Dawson, Brandy Vela, Amanda Todd, Tyler Clementi, Megan Meier, and Callum Moody-Chapman committed suicide after being trolled and/or receiving abuse online. Whilst cyber-bullying and trolling does not necessarily cause suicide, research has illustrated that there is an association between the two (Messias, Kindrick and Castro, 2014). Other negative experiences of victims of trolling and online abuse include depression and anxiety (Selkie et al., 2016). Victims of cyberbullying are also more likely to be involved in alcohol and drug misuse (Hinduja and Patchin, 2014). Additionally, victims, and even observers of online abuse have stopped using social media altogether as a result of trolling (Tiku and Newton, 2015), including Leslie Jones following the onslaught of racist comments. Whilst trolling can be positively marked, these negative and devastating effects have led to various suggestions for regulation and moderation.

Moderating online spaces is a key challenge. There are two kinds of moderation: distributed social moderation and machine-learning based algorithms. Distributed social moderation involves community members reporting on or voting on cases of trolling and then the site can take action, whereas machine-learning based algorithms involve training classifiers to detect trolling and other forms of abusive language based on datasets of previously detected abusive posts (Chandrasekharan et al., 2017). Over the years, Twitter has been scrutinised for the way in which it has dealt with trolling and abuse. Hicks and Gasca (2019) note that Twitter has previously tended to only review potentially abusive tweets that were reported to them by Twitter users. However, more recently, especially in 2019, Twitter has begun to be more proactive by using technology, which has proven to be successful in detecting a proportion of abusive behaviour, hateful conduct, threats and violence (Hicks and Gasca, 2019).

1.3. Aims and Contributions of this Dissertation

Importantly, we know that trolling is used to label a variety of different behaviours with different communicative goals, from the spread of misinformation and hyperbole to targeted abuse and malicious attacks. When so much variation exists, the question arises about whether the diversity of behaviours encapsulated in the usage of the term is actually reflected in the language across numerous instances of trolling. Nevertheless, given that trolling is a linguistic phenomenon, little is known about how trolling varies linguistically from one to the next and what the major linguistic repertoires and communicative functions are. This dissertation therefore aims to collect a corpus of trolling tweets and subsequently provide a thorough linguistic description of those trolling tweets by identifying and describing the major patterns of linguistic variation and the major communicative functions across them.

The challenge in collecting such a corpus lies in being able to detect trolling in the first place, especially considering its diverse and deceptive nature. Pragmatically, this dissertation identifies and collects trolling instances based on other people’s perceptions and accusations. Essentially, if a tweet is labelled or accused to be trolling, then it was collected. Having collected the corpus of trolling tweets, the dissertation analyses and describes the major communicative functions and patterns of linguistic variation across them. One method commonly used to identify the major communicative functions of a language variety is Multi-Dimensional Analysis (MDA) (Biber, 1988). MDA is based on the assumption that frequent patterns of co-occurring linguistic features tend to suggest at least one shared communicative function (Biber, 1988). In other words, if many texts share similar frequencies of particular linguistic features, then it is likely that they will share at least one communicative function. Based on this assumption, MDA is aimed at identifying the major patterns of linguistic co-occurrence across a corpus of texts, and then these patterns are interpreted for their underlying communicative function.

There are, however, two problems with applying MDA to tweets. The first problem is that the tagger used to identify the linguistic features in each text for the analysis is not well suited to tweets. The second problem is that tweets are characteristically short texts and rarely exceed 40 words. For the most part, MDA requires texts to be over 500 (Passonneau et al., 2014) or 1000 words long (Biber, 1993) – far longer than tweets. Consequently, this dissertation introduces a new tagger and also a modified short text version of MDA, and applies this to the corpus of trolling tweets.

Whilst this analysis identifies the dominant linguistic co-occurrence patterns across the corpus of Twitter trolling, it is not clear whether these major patterns of linguistic variation are unique to Twitter trolling or whether they are just general patterns of tweets more generally. Given that Twitter trolling is situated in the context of Twitter, trolling tweets could just be drawing on the major linguistic repertoires of general tweets. It is therefore important that the major patterns of linguistic variation of trolling tweets be compared with the major patterns of linguistic variation of general tweets. Nevertheless, there has not yet been an analysis of the major patterns of linguistic variation across general Twitter with which to compare the results of the Twitter trolling analysis. In particular, little is known about the major linguistic repertoires and communicative functions of general tweets. Consequently, this dissertation also collects a corpus of general English tweets and runs a separate MDA on this corpus of general tweets in order to identify and describe the major patterns of linguistic variation of general Twitter. The results of this analysis are described in this dissertation prior to the results of the MDA of Twitter trolling. Importantly, these results serve as a baseline to which the MDA of trolling tweets is subsequently compared. Specifically, the separate MDA of general Twitter enables the comparison of the major communicative functions of general English tweets and trolling tweets.

Because trolling tweets do not exist outside of Twitter, but are situated in the context of Twitter, it is interesting to observe whereabouts trolling tweets are positioned with respect to general tweets. Specifically, whilst we know that trolling tweets are a kind of tweet, we know very little about which tweets are most similar to trolling tweets and what kinds of linguistic repertoires of general tweets do trolling tweets align with the most. Consequently, in the final analysis chapter, this dissertation positions trolling tweets along the dimensions of linguistic variation of general tweets in order to quantitatively and systematically compare the tendencies of trolling tweets and general tweets with respect to the major communicative functions of general English tweets. Overall, these analyses enable the thorough linguistic description of Twitter trolling, not only in terms of its major communicative functions and patterns of linguistic variation, but also with respect to the major communicative functions and patterns of linguistic variation of general Twitter.

1.4. Outline of the Dissertation

Following this introductory chapter, this dissertation is organised into 7 subsequent chapters. Chapter 2 offers a review of previous literature and in light of this presents the major research questions that this dissertation attempts to answer. Chapter 3 describes the method for collecting the data and Chapter 4 describes the method used to answer these questions. Chapter 5 presents the results of the MDA of general English Twitter. Chapter 6 presents the results of the MDA of Twitter trolling and compares these results to the results from Chapter 5. Chapter 7 presents the comparison of trolling tweets to general English tweet along the dimensions of general English Twitter. Finally, Chapter 8 revisits each research question, providing a summary of each study and a conclusion. Additionally, Chapter 8 discusses the contributions of this dissertation and avenues for future research.

2. Literature Review

This dissertation aims to identify and describe how trolling tweets vary linguistically not only from one trolling tweet to the next but also in relation to general tweets. This review is therefore aimed at bringing together sociolinguistic research on language variation on social media, especially Twitter, as well as research on trolling.

2.1. Sociolinguistic Variation on Social Media

There are several ways within the linguistics tradition to analyse linguistic variation, including the variationist approach, systemic functional linguistics, and the corpus-based text linguistic approach (see Lee, 2001; Bawarshi and Reiff, 2010; Rickford and Eckert, 2001). The central tenet to all approaches to language variation is that linguistic variation follows patterns that are regular and probabilistic, and these can be explained through extra-linguistic factors, such as demographic variables, the situational characteristics or register, and the communicative function (Bohmann, 2017). Many of the approaches differ in respect to their object of analysis, ranging from individual linguistic variables or texts. Thus, the approach that the researcher takes to examine language variation influences the data required, and by extension, the method for collecting data.

Traditionally, under the variationist approach, language variation and change have been described by predominantly collecting data using questionnaires and surveys, and interviews (Weinreich, Labov, and Herzog 1968; Trudgill 1974; Milroy and Milroy 1985; Eckert 1989; Milroy and Gordon 2003; Tagliamonte 2006), whereas researchers using SFL and the corpus-based text linguistic approach use collections of texts from particular registers (e.g. fiction, academic essays). More recently, social media data is being used across all approaches for examining language variation and change.

For example, research has sought to understand the new registers across social media, and have therefore used social media data to describe patterns of register variation, that is, identifying linguistic similarities and differences across social media and other online and offline registers (Sardinha, 2014; Sardinha, 2018; Titak and Roberson, 2013; Friginal, Waugh and Titak, 2018). Specifically, Sardinha (2018) collected a corpus of various online registers including emails, tweets and Facebook posts and applied Biber's (1988) corpus-based text linguistic approach, called Multi-Dimensional Analysis, revealing 3 major dimensions of register variation. The first dimension opposes registers that are more involved and interactive with registers that are more informational. The second and third dimension are associated with expressing stance (evidential and affect). The fact that both of these dimensions are stance related marks the importance of expressing stance across social media.

In addition to describing patterns of register variation, social media has been used to understand regional linguistic variation, that is, identifying patterns of linguistic variation across social media and mapping them to particular dialect regions (Grieve, 2016; Huang et al., 2016 for U.S; Nguyen and Eisenstein, 2017 for Netherlands; Donoso and Sanchez, 2017 for Spain), gender and age variation, that is, identifying patterns of linguistic variation according to particular age and gender groups (Pavalanathan and Eisenstein, 2015), language choice, that is, identifying which factors influence the language choice of multilingual users on Twitter (Eleta and Golbeck, 2014), lexical emergence, that is, identifying emerging word forms in large corpora of time-stamped Tweets (Grieve et al., 2017a), lexical innovation, that is, tracking the origin and spread of new words on Twitter (Grieve et al., 2018b), as well as understanding and describing patterns of language change over time, such as patterns of increasing and decreasing words on Twitter (Nini et al., 2017). Moreover, social media data has been used to understand the structural and social role of language change and linguistic dissemination, that is the distribution of a new word or construction across lexical contexts (Stewart and Eisenstein, 2018).

Social media has also been used to test the robustness of various sociolinguistic assumptions and theories. For instance, based on the assumption of linguistic homophily (whereby socially connected individuals use language more similarly than those who are not) (Yang and Eisenstein, 2017), and the claims that part-of-speech taggers' performance are hindered by stylistic diversity, Balusu, Merghani and Eisenstein (2018) compared the error rates of part-of-speech taggers with network structure on social media, finding that in some parts of the network there was higher accuracy than in other parts, suggesting that there is support for the notion of linguistic homophily.

Additionally, linguistic theories such as Labov and Waletzksy's (1967) narrative theory (e.g. Dayter, 2015), Giles' (1984) communication accommodation theory (e.g. Danescu-Niculescu-Mizil, Gamon and Dumais, 2011), and Bell’s (1984) theory of audience design have been tested for their robustness using social media data. For example, with the general move towards the sociolinguistic notion that identity is performed and varies according to the context, Bamman, Eisenstein and Schnoebelen (2014) used cluster analysis on an individual’s use of particularly distinctive words, as well as information about their social network (i.e. their reciprocal followers) in a quantitative analysis of linguistic style (here referring to lexical variation) and the sociolinguistic notion of gender identity. They found that the clusters often relate to particular topics with many also being associated to a particular gender, although sometimes the lexis that the particular gender employed did not match the overall population-level language statistics. They explain this finding with information from the social network. They found that these ‘outliers’ have significantly fewer same-gender followers, and consequently suggest that both communication accommodation (Giles, Coupland and Coupland, 1991) and audience design (Bell, 1984) are at work here, whereby the authors of the posts are employing particular linguistic resources to position themselves in line with their social network connections - that is, performing a more masculine identity, evidenced through lexical choice, to a male-dominated social network.

In addition, Pavalanathan and Eisenstein (2015) suggest that they find support for audience design in Twitter data, as they discover that individuals are more likely to Tweet nonstandard lexical variables when the audience of the message is smaller than when the audience is large, evidenced by the use of hashtags, which can function to broadcast the message to particular audiences beyond one's followers (Pavalanathan and Eisenstein, 2015). Moreover, based on the assumption that people's social identity is reflected in their language use, Shoemark et al. (2017) hypothesised that those in favour of Scottish independence (during the 2014 Scottish independence referendum) would be more likely to use distinctively Scottish terms in their social media posts than those who were anti-independence. To test this hypothesis, Shoemark et al. (2017) conducted a large-scale study of sociolinguistic variation in tweets sent in the UK. Specifically, they investigated the use of distinctively Scottish lexis and their Standard English counterparts amongst pro- and anti-independence hashtag users, finding that such Scottish items were used at a higher rate in the tweets of pro-independence users than in those by anti-independence users. Despite this, they found that the tweets containing the referendum-related hashtags contained far fewer distinctively Scottish words than the individuals' general tweets, which they suggest supports what Pavalanathan and Eisenstein (2015) found in their study, which is that Twitter users tend to use fewer non-standard and local variants when the expected audience is larger, as would be the case when using a hashtag.

Other researchers have used social media data to test social network theories, such as complex contagion, whereby the likelihood of certain forms being adopted increases with the amount of exposures (Goel et al., 2016), as well as understanding the creation of social ties, testing various predictors, such as common friends and common interests (Hours, Fleury and Karsai, 2016). For example, Goel et al. (2016) reveal that complex contagion is observed with phonetic spellings and abbreviations (e.g. lol, lmao). Hours, Fleury and Karsai (2016) quantify the common interest that two users may share by measuring hashtag similarity, which they describe as the distance between the sets of hashtags that the user tweeted, as well as common friends, finding that these measures are correlated with connected people.

An extension of distributional semantics, which is based on the theory that semantically similar words occur within the same textual contexts, has also been tested using Tweets. Cocos and Callison-burch (2017) extended this theory to investigate whether the same is true for words that occur within the same geospatial context. Specifically, they investigated the extent to which semantically similar words occur within the same geospatial context by supplementing geotagged Tweets with more general categories of places using Google Places and OpenStreetMap (e.g. bar, residential). They found that "people are more likely to tweet about something they *love* from a bar than from home, but vice versa for something they *hate*" (Cocos and Callison-burch, 2017: 99). Overall, they found that textual content was far more informative than geospatial context, although the geospatial context still encoded information about semantic relatedness, which they suggest can be used to complement the semantic information retrieved from textual context as part of a multimodal model.

In addition to testing various sociolinguistic theories and assumptions and understanding and describing patterns of sociolinguistic variation, researchers within the field of computer science and natural language processing (NLP) have sought to develop tools and models that use the text and other features in order to predict particular social variables, such as age (Schler et al. 2006; Rosenthal and McKeown 2011; Nguyen et al 2013; Liao et al. 2014), gender (Koppel, Argamon and Shimoni, 2002; Herring and Paollillo, 2006; Schler et al. 2006; Mukherjee and Liu, 2010; Rao et al. 2010; Burger et al. 2011; Bamman, Eisenstein and Schnoebelen 2014), race (Eisenstein, Smith, Xing 2011), geography (Eisenstein et al. 2010; Rahimi, Cohn and Baldwin, 2017), and even political affiliation (Conover et al. 2011; Cohen and Ruths, 2013; Tatman et al., 2017). Not only is this social information used to understand sociolinguistic variation and change, but it is also used to improve the performance of existing NLP tools (Hovy, 2015). Other research has sought to identify language variation predictors of particular social relationships, such as when there is a power imbalance (Danescu-Niculescu-Mizil et al. 2012; Gilbert 2012; Prabhakaran, Rambow, and Diab 2012), or for politeness (Danescu-Niculescu-Mizil et al. 2013).

Many of these studies have required large corpora of texts that are densely sampled over periods of time. This is because most lexical words are actually very rare, with the majority of words occurring less than once per million words (Grieve et al., 2017a). The rarity of most lexical forms is referred to in the literature as Zipf’s law (1949) or as the “Long-Tail Problem” (Eisenstein et al., 2014), which describes the plot of the frequency of lexical items in corpora with only a very few amount of words being used very frequently and then this number dramatically decreases with the majority of words appearing relatively few times. Additionally, some new forms have been shown to rise extremely quickly in frequency in a short amount of time, meaning that the corpora not only need to be large, but also densely sampled over time (Grieve et al., 2017a). Social media data, especially Twitter, which is used in the majority of the studies above, has therefore facilitated such investigations because large amounts of data can be collected with relative ease. Moreover, quantitative and more robust methods have also been introduced into the field as a result of such large corpora.

Social media data has also been used in social semiotic research to identify the interrelation between particular platforms and the technology and the sign-maker’s expression. Introduced to linguistics by Halliday (1978), social semiotics is, for the most part, a multimodal approach, which is centred on the process of sign-making. Social semiotics seeks to understand how the social setting influences ways of communicating. In essence, it suggests that communicators have the choice to represent their interests, but do so by selecting from the repertoire of resources that are made available in the particular context. These contexts and the resources have affordances, including limitations to what signs can be made, as well as possible uses for how certain things can be represented. The context may also foreground or prefer particular meanings, and may also simultaneously background and disprefer particular options for sign making (Adami, 2018). For many this raises important questions about the extent to which “the technological and social forces drive the sign-makers preferred selections of resources” (Adami, 2018: 602). Social semiotic studies of social media include examining various texts and practices on platforms, such as practices of self-expression and identity construction on weblog publishing platforms (Adami, 2018) and Facebook (Bouvier, 2012), as well as hashtagging and retweeting on Twitter (Zappavigna, 2018), and selfies on Instagram (Zappavigna, 2016). These studies reveal how the technology constrains and affords particular meanings and patterns of variation.

This section demonstrates the benefits of using social media data in linguistic research: not only are we able to collect vast amounts of more natural data and ask different and arguably more general questions about language variation and change than ever before, but we are also able to answer them in a more efficient, systematic and methodologically robust way. Additionally, we are able to test the robustness of existing sociolinguistic theories with social media data and explore the influence of technology and platforms on particular social practices. This body of research importantly also shows that language varies, not only across social media platforms and registers, but also within platforms, according to numerous extra-linguistic factors.

2.2. Twitter

Twitter is one popular social media platform that has been used considerably in numerous research fields and in many of the studies that were mentioned in the previous section. As the focus of this research, this section aims to describe Twitter in terms of its infrastructure, its uses, its user base and its rules in order to situate it as a semiotic technology, which has particular affordances that influence the social practices and language on Twitter. Moreover, this section reviews previous research on the language of Twitter as an online register in relation to other varieties of language. Finally, Twitter is positioned as a platform that is used for a variety of purposes and plays host to and encourages an abundance of linguistic variation, including more sinister varieties of language, although the same could be said for other forms of social media.

2.2.1 Infrastructure, User Base, Uses and Rules

Twitter was created in 2006 and became popular quickly. Although there are debates about what Twitter was intended for and what Twitter is, it is popularly characterised as a social networking micro-blogging service, which allows users to post messages called ‘tweets’, to a network of associates, deemed ‘followers’. Each user chooses a username preceded by the ‘@’ symbol (e.g. @issy_clarke1). They also have a profile (see Figure 1), which compiles all their previous tweets. Users can select an image for their profile and an image for the background of their profile. Twitter also provides a short text box, called a ‘bio’, which tends to invite short biographical descriptions. These affordances enable purposeful identity construction.

In the design of Twitter (see Zappavigna, 2018), tweets were originally constrained to 140 characters long so that they could be sent using SMS, which were, on most phones, restricted to 160 characters, permitting 20 characters for usernames. But since November 2017, the character restriction per Tweet has doubled to 280 characters. Recent research examining the effect of this shift on tweet success and language style has found that users wrote more tersely and included more abbreviations and contracted forms and fewer definite articles when there was a length constraint (Gligorić, Anderson and West, 2018). Thus, the technological affordance of tweet length has been shown to influence style. In addition to SMS, tweets can be posted through the web, applications and third-party clients on phones, computers and other devices, enabling users to tweet whilst on-the-go.

The user base of Twitter is defined in various ways. Twitter, for instance, refers to active users, where ‘active’ refers to individuals who were logged in or who were authenticated and who accessed Twitter via the website or third party applications (Twitter Inc., 2018). This suggests that users do not have to tweet to be active, but instead can lurk and observe content on Twitter. Based on this definition of active users, for the first quarter of 2019, Twitter revealed that it had 330 million monthly active users (Clement, 2019).

The user base is also defined demographically. There are 69 million Twitter users in the U.S., which equates to approximately 21 percent of all Twitter accounts. Thus, 79 percent of Twitter accounts are based outside of the U.S (Twitter Inc., 2018) in over 150 countries. After the U.S., the top three countries by user count are Brazil (27.7 million), Japan (25.9 million) and Mexico (23.5 million) (eMarketer.com, 2016). The ages of Twitter users vary, but 38 percent of users fall between the ages of 18 to 29, 26 percent of users are between 30 and 49 years old, 17 percent of users are between 50 and 64, whilst 7 percent are 65 and above (Pew Research Center, 2019). In addition to this demographic information, the user base of Twitter is also defined by social status, where celebrities, politicians and persons of public interest are authenticated and assigned a blue tick.

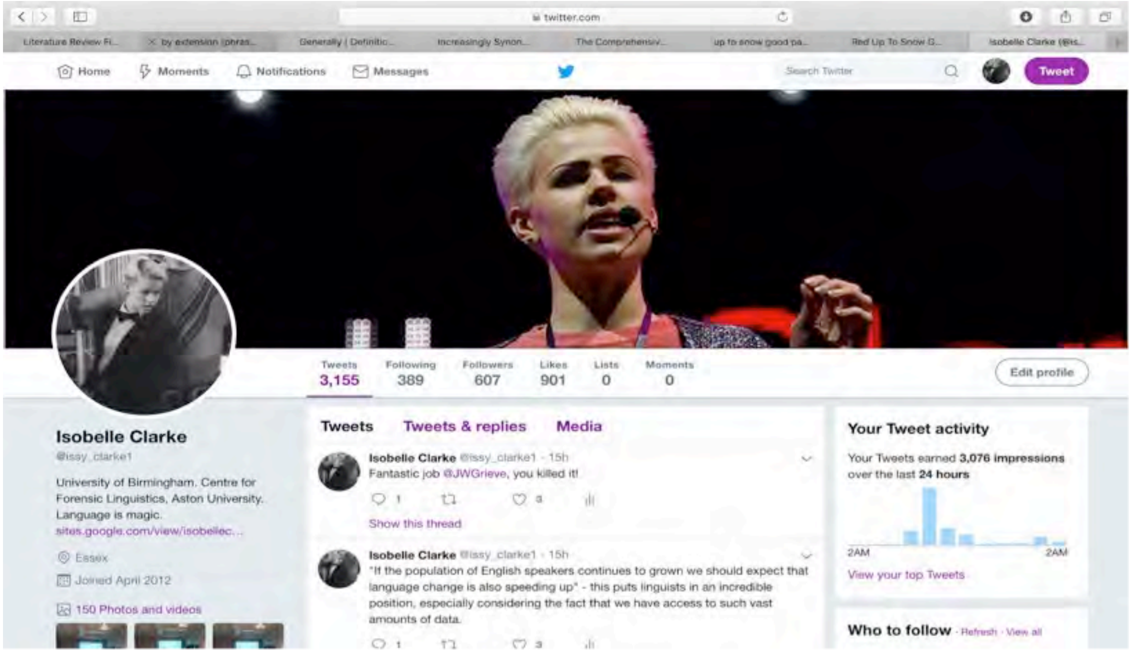
The user base has also been defined by the ways in which they use Twitter, including following and posting habits. Twitter is based on ‘following’, which can be thought of as similar to subscribing to someone’s updates.

Following is non-reciprocal; that is, a user does not have to follow someone back if they follow them, leading to the formation of complex follower networks of unidirectional and bidirectional connections (Weller et al., 2014). For example, my Twitter page is in Figure 1 and it shows in big and bold on the header that I am following 389 people and have 607 followers. Such connections are articulated - one is able to click on the following or followers tab (see Figure 1) to observe a user's audience and the individuals that they follow (boyd and Ellison, 2007; Schmidt, 2014). Krishnamurthy et al. (2008) examined the ratio between following and followers, as well as tweeting habits of over 100,000 users. They found three distinct groups of tweeters: (1) broadcasters, those who tweet often and have a larger amount of followers in comparison to the amount they follow; (2) acquaintances, those who have a relatively equal amount of followers in respect to the amount of people they follow; and (3) miscreants or evangelists, those who have few followers but follow several people. Java et al. (2007) found similar categories of Twitter users, but labelled the users with the most followers who post news as 'information sources', whilst those who follow several users but rarely post were labelled as 'information seekers'.

Previous research has described the influence of following and follower networks on the communicative practices on Twitter. Schmidt (2014), for example, argues that deciding whether to tweet will be based on the user's perception of their audience. Over time, however, especially as the communicative practice of tweeting becomes more routinised, the extent of scrutiny that each tweet receives might reduce, and rather the user will design their tweet and assess its appropriateness according to their *intended* *audience* (Schmidt, 2011) or *imagined audience* (Litt, 2012), and even a particular group within that audience (Schmidt, 2014). For example, my audience on Twitter could be divided into academics and non-academics; however, the majority of my tweets are designed for the academic community and assessed for their appropriateness according to this community's standards. Schmidt (2014: 11) describes this as "privacy management".

Page (2012) similarly suggests that certain Twitter practices have been influenced by the user's perception of their audience and the notion of following. Specifically, Page (2012) notes that the non-reciprocity of following has enabled some individuals to have millions of followers, especially celebrities, prominent figures and businesses. As a result, the size of a follower list is often taken as a sign of status and influence. Jack Dorsey (2019) the CEO of Twitter also attributed the importance assigned to follower size to the architecture of the platform because 'following' and 'followers' are presented in big and bold (see Figure 1), which ultimately incentivises gaining followers. Regardless of why, because of the importance attributed to follower size, Page (2012) and Dorsey (2019) suggest that there is a need and incentive amongst Twitter users to have and gain more followers. It has been suggested that this need to have more followers has influenced practices of promoting, self-branding, micro-celebrity (Page, 2012), and more oppositional, provocative and uncivil content (Anderson, 2019), which will be discussed in section 2.2.3. Thus, the technological affordance of following and the meaning ascribed to follower size afforded by the design of the platform has influenced social practices.

Figure 1: Isobelle Clarke's Twitter profile

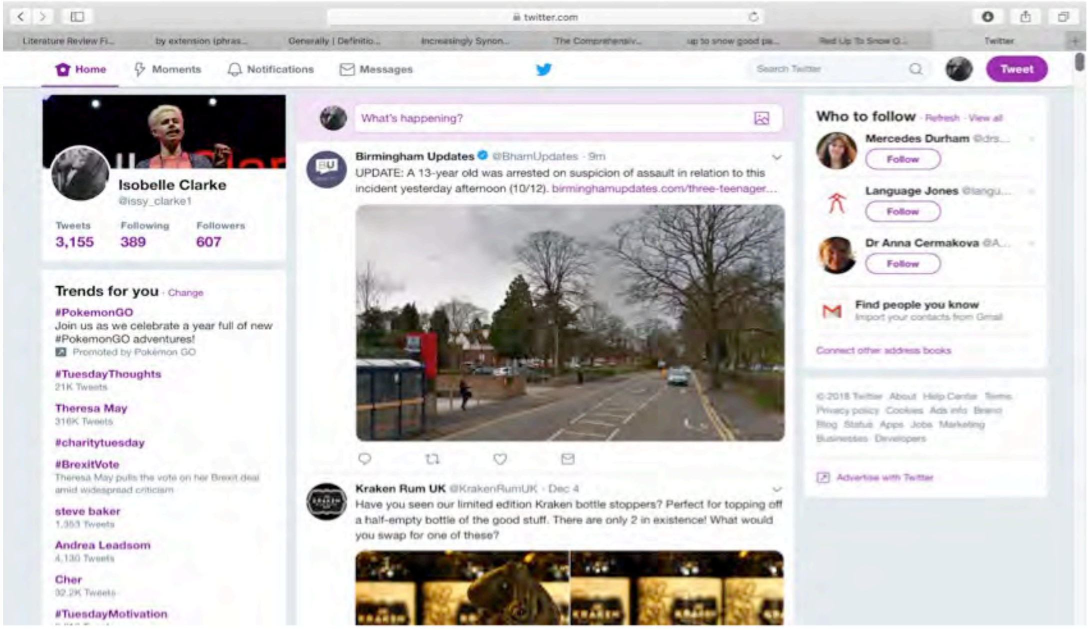


Over the years, Twitter has changed the appearance and structure of the service, adding new features, often as a result of user-generated conventions (Weller et al., 2014). For example, users began using the hashtag before the most important word in the tweet to tag them, and, as a result of its prominence, it was later incorporated in the service as a feature to signal and identify trends (Weller et al., 2014; see Halavais, 2014). Additionally, users were interacting with other users by including the user's username, preceded by the '@' symbol, which is called 'mentioning'. For example, Figure 1 shows that I have mentioned @JWGrieve in the tweet beginning with 'Fantastic job…". As a result of its prominence, '@' mentioning was incorporated into the platform service, encouraging conversation and collaboration on Twitter (Honeycutt and Herring, 2009).

In the most recent form of Twitter (July 2019), a user posts a tweet in the box that has a prompt, which says: “What’s happening?”, which disappears once the user clicks on the box and begins typing their tweet (see Figure 2). This prompt previously said: “What are you doing?”. Zappavigna (2015) suggests that this change in prompt influenced users into posting tweets that not only concerned personal activities, but also tweets on topical events and a wide range of phenomena, which made Twitter into a platform for not only self-expression, but also news distribution.

Twitter is frequently thought of as a medium for instantaneous news distribution. Sveningsson (2014), for example, interviewed 26 Swedish young people and found that social media, especially Twitter, is their first point of reference to obtain news. For example, Twitter is often used to obtain and contribute to real-time information on events, such as disasters like the Pakistani floods in 2010 (Murthy and Longwell, 2013), protests, such as the black lives matter movement (Freelon, McIlwain and Clark, 2016) and “Arab Spring” uprisings in 2011 (Bruns et al., 2013), riots, such as the UK riots in 2011 (Proctor et al., 2013), televised sporting events (Lim, Hwang, Kim and Biocca, 2015; Neeley-Cohen, 2016), political debates and elections (Gottfried, Hardy, Holbert, Winneg and Jamieson, 2017; Bastos et al., 2013), and other entertainment and news events, such as the Oscars, the death of Michael Jackson, and the royal weddings (Weller et al., 2014). These events can be on a global level, such as those just described, as well as on a more local level (Weller et al., 2014), such as the Birmingham Updates account on Figure 2, which reports on local news and events in Birmingham, UK.

Figure 2: Isobelle Clarke's Twitter Timeline



Once a tweet has been sent, the user's Followers can see the tweet on their Timeline. Figure 2 is a screenshot of my Timeline. The basic concept of following means that each user's timeline is comprised of all the tweets of the people that one follows. Originally, the tweets of the people that users follow were presented in reverse-chronological order (Schmidt, 2014). However, in 2016 Twitter launched an algorithm, which decides which tweets the user gets to see. Specifically, it is aimed at ensuring that users get to see the tweets from the people that they interact with the most and that they get to observe the most popular tweets out of the people they follow (Oremus, 2017). According to Oremus (2017), this algorithm was a consequence of a shift in the use of Twitter from status-updates to a news platform, and the sheer effort required to find important tweets when tweets were organised reverse-chronologically:

> *The reverse-chronological timeline stemmed from the site's origins as a way to blast brief, real-time "status updates" via text message to friends and acquaintances. But over the years Twitter morphed into something more like a public platform for news, opinions, jokes. As the user base and its follow lists grew, the chronological feed's limitations became clear. You'd log in and find yourself thrust into the middle of dozens of unrelated, often insider-y conversations, and the good stuff required tedious scrolling to unearth. For the ordinary internet user, it simply wasn't worth the trouble.*

This suggests the stream of information that appears on one's timeline is filtered according to the social connections that one has made on Twitter (Schmidt, 2014), and also according to the user's favourites and the popularity of the tweets of the user's connections.

Tweets from people outside of their social connections can also appear in a user's timeline, as a result of retweeting. Retweeting is where a tweet from a user is reposted by one of their followers and is shared to all of the reposter's followers. Retweeting occurs frequently for many different reasons, including to distribute news quickly to one's followers, especially if one's followers do not follow the particular account, to signal that one shares the views expressed in the tweet, to raise awareness of it's content and others (see boyd et al., 2010). Retweeting means that a tweet can reach a network of people beyond one's own followers, and even expectations. For example, Justine Sacco posted a racist 'joke' prior to setting off on a plane journey to Africa (Pilkington, 2013). During that plane ride, her tweet had been retweeted over 2,000 times; she had become a trending topic, and was met with hostility from the Twittersphere. As a result of her comment, she was fired from her job.

In addition to merely forwarding the message, users can also add a comment alongside the retweeted message, enabling a stance and/or evaluation to be appended to the message (Page, 2012; Zappavigna, 2018). Thus, retweeting also “contributes to a conversational ecology in which conversations are composed of a public interplay of voices that give rise to an emotional sense of shared conversational context (boyd et al., 2010: 1). In order to retweet, users originally copied and pasted the original tweet and signalled that it was a retweet with ‘RT @username:’. However, due to its prominence on the platform, the content creators integrated this into the design of the platform. Specifically, the symbol appears underneath all tweets and can be clicked on by other users if one’s account is public to retweet the particular post. The tweets of private users cannot be retweeted and the symbol is dimmed and unclickable. Twitter communication is therefore based on textual references made explicit via particular software affordances, such as the ‘@’ symbol, retweeting and hashtags (described below), and these provide visible and navigable communicative references to other Twitter users (Schmidt, 2014).

In addition to following accounts, Twitter provides a search tool, whereby users have the opportunity to search for particular words, phrases or people (see top right of Figure 1 and Figure 2), so that they can follow their interests and hear what people are talking about. Users can create, contribute to and interact with particular themes and content by using the hashtag symbol ‘#’ before a particular string of letters, words, or phrases. Hashtags are also searchable and have been found to occur to reference the target of evaluation (Zappavigna, 2011). Hashtags afford the formation of relations between users and texts on Twitter (Zappavigna, 2015; Schmidt, 2014). This is because of their searchability, which ultimately connects the tweets of users who use the same hashtag, despite the fact that they may not even follow each other. Given their searchability, users are able to search for and observe tweets that are tagged in a specific way and contribute to the topic, expressing similar sentiment, and by extension, showing affiliation (Zappavigna, 2018). Hashtags are not only used to reference the target of tweets, but are also used for a wide variety of communicative functions, including gaining visibility and evaluating (Page, 2012), playing games, meta-commentary, and for referencing popular culture and memes (Wikström, 2014), potentially in order to demonstrate cultural knowledge. Overall, this means that the stream of information that a user may encounter is also filtered according to particular phrases or keywords that one is interested in or uses.

If several people use particular hashtags in a short time frame, they can become popular and categorised as ‘trending’. Trending topics throughout the Twittersphere are displayed on the left hand side of a user’s home Timeline (see Figure 2). PokemonGO and Theresa May are trending topics on Figure 2. Tweets can also include URLs to other sites and other types of multimedia, such as images, videos and animations. The URL can function to support the text (e.g. He only went and did it [picture of a hand with a ring on the engagement finger]); it can also be the object of the text’s content (e.g. *This is crazy* [video of a magic trick]). URLs can also provide additional content, especially considering the length restrictions on tweets, enabling the content of tweets to go further. For example, in Figure 2, the tweet from Birmingham Updates contains a URL to a news article. URLs can be quite long and with the 280-character restriction, this can be problematic, so Twitter provides a tool to shorten them to 23 characters in order to save space. A tweet can also be ‘liked’ by clicking on the heart and once a user has liked a tweet, these tweets are stored in their Likes tab on their homepage (see Figure 1). This indicates that Twitter essentially has a rewards system in place, where tweets are rewarded with likes.

Overall, Schmidt (2014: 6) states that these particular “affordances of Twitter as a software service, together with the social and textual affordances articulated in ongoing use, form a communicative space which is partly stable (e.g., the connection between followers and followees) and partly highly dynamic (e.g., the tweets using a popular hashtag)”. In other words, Schmidt (2014) argues that the communicative space on Twitter is structured via the technological affordances and features of the platform and the social and textual relations that are made explicit, although the way in which these features are used can vary. This is in line with social semiotics as they position the semiotic resources as having meaning potential, and the different meanings will be based on the user’s interest in the particular communicative context (van Leeuwen, 2005; Kress, 2010). Nevertheless, users of a platform are constrained by the affordances of the technology, including what semiotic resources are made available and accessible. For instance, the technology constrains the length of tweets, which means that longer messages either have to be condensed or spread out across numerous tweets. Additionally, Twitter is not predicated on instant messaging, and therefore, there is no feature which reveals whether someone is in the process of tweeting, as there is with Whatsapp and Facebook messaging.

In addition to constraints imposed by the technology, Twitter also restricts particular behaviours in its rules. These rules fall into three categories: content boundaries and use of Twitter, abusive behaviour, and spam and security (The Twitter Rules, 2018). With respect to content boundaries and use of Twitter, the rules concern using Twitter for illegal means, distributing hacked materials, intellectual property (e.g. trademark and copyright), misusing the Twitter badge or using it without permission, advertising without permission, manipulating trends and uploading graphic content (e.g. pornographic and violent media). Twitter also prohibits username squatting and selling, which involves creating several accounts in order to prevent other people from having those usernames with the intention to then sell these to those users. The rules within the 'spam and security' category concern sending mass or repeated invitations or malicious content intended to damage property or obtain confidential information, as well as creating fake accounts. Twitter's rules on abusive behaviour begins with the following statement:

> *We believe in freedom of expression and open dialogue, but that means little as an underlying philosophy if voices are silenced because people are afraid to speak up. In order to ensure that people feel safe expressing diverse opinions and beliefs, we prohibit behavior that*

> *crosses the line into abuse, including behavior that harasses, intimidates, or uses fear to silence another user's voice.*
> *Context matters when evaluating for abusive behavior and determining appropriate enforcement actions. Factors we may take into consideration include, but are not limited to whether:*
>
> - *the behavior is targeted at an individual or group of people;*
> - *the report has been filed by the target of the abuse or a bystander;*
> - *the behavior is newsworthy and in the legitimate public interest.* (Twitter, 2018)

Twitter also specifically prohibits publicising suicide, self-harm, violence and child sexual exploitation. Twitter forbids abuse that harasses, intimidates or silences. Twitter prohibits unwanted sexual advances, including sending unwanted sexual content, and hateful conduct such as threatening, harassing or promoting violence on individuals because of their race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. Twitter also forbids publishing private and confidential information (i.e. doxing), and intimate media (e.g. revenge porn), or threats to do so. Finally, Twitter prohibits impersonation and having multiple accounts for similar purposes, especially if the initial account has been suspended or banned. Users who engage in any of the following activities are likely to have their account suspended either temporarily or permanently and/or their information passed on to law enforcement.

Twitter has in place algorithms and tools used to monitor suspicious activity, as well as to hide sensitive material. Nevertheless, with over 500 million tweets per day, it is difficult to remove or censor all posts which break these rules. Moreover, intentional rule breakers have found covert ways to bypass these algorithms, such as through creative spelling and avoiding profanity (Hine et al., 2017). They therefore also provide a tool for users to report activities deemed to go against the rules. However, despite such moderation, Twitter has received considerable criticism for the way in which it deals with ‘rule breakers’ and/or trolling and other forms of abusive language.

One of the reasons for this is because Twitter is a constellation of millions of communities, each with their own set of norms and expectations. Such expectations about how individuals should use Twitter, i.e. what is normal Twitter use, are not explicitly laid out (Schmidt, 2014). Rather, these rules might be made apparent when certain behaviour is deemed inappropriate, or if there are conflicts, misunderstandings or communication breakdown (Schmidt, 2014). Importantly, these rules will be specific to the particular communities in which one resides. Thus, in addition to the service enforcing rules, there are also other rules and regulations constructed by the users of Twitter. For example, it is generally considered to be improper to retweet someone’s tweet if that person has set their account, and by extension their tweets, to private. Twitter does not allow individuals to retweet a person whose account is private by clicking the retweet symbol. Users, nevertheless, are able to retweet, although this requires a bit more effort, by using ‘RT’ and copying and pasting the tweet; however this can result in criticism.

Overall, this section has attempted to position Twitter as a semiotic technology by exploring the relationship between the platform, the technology, the users, the texts and the social practices on Twitter. The various uses and norms of Twitter have been co-created over time by Twitter specifically, as well as users and communities of users through their practices of use (Weller et al., 2014). By using the service, users shape the service, and this has led to new semiotic resources, tools, phenomena and forms of communication in participatory culture (Havalais, 2014; Weller et al., 2014), such as hashtags and trending information, as well as live-tweeting during major news and televised events. This review reveals that Twitter has preselected technological features and affordances. These technological affordances (e.g. character restrictions, follower-followee networks, URL shortening), the social and textual relations (e.g. hashtagging, mentioning, and retweeting), and the shared rules and community specific expectations all influence the language of Twitter.

2.2.2. The Language of Twitter

Twitter is used for a variety of purposes, such as for disseminating information on major news events, for personal expression and interaction (Honeycutt and Herring, 2009), as well as for recording one’s thoughts and narrating one’s everyday activities (Weller et al., 2014), among many others, often according to personal preference (Schmidt, 2014). Nevertheless, the full extent of the linguistic variation across Twitter has yet to be examined.

Previous research using the corpus-based text linguistic approach, specifically Biber's (1988) MDA, has explored how Twitter as a variety of language compares to the major patterns of linguistic variation of other varieties of online language (Titak and Roberson, 2013; Sardinha, 2018), other varieties of social media (Friginal, Waugh and Titak, 2018), other registers (Passonneau et al., 2014), and pre-internet registers (Sardinha, 2014); nevertheless, the majority of these studies ignore the variation that occurs from one tweet to the next. MDA is based on the notion of linguistic co-occurrence (Ervin-Tripp, 1972; Hymes, 1974; Brown and Fraser, 1979) - patterns of co-occurring linguistic features are not random, but rather there is an underlying cause for the linguistic co-occurrence patterns, which is suggested to be a shared underlying communicative function and/or shared situation. MDA tends to assume that there is not one parameter of variation that explains all language in a particular domain, but rather multiple dimensions of linguistic variation will be operating in any discourse domain. These analyses therefore involve trying to define the overall dimensions of variation within a particular domain (Biber, 1988; Biber 1989), so that similarities and differences amongst registers and particular texts can be described. For example, in previous analyses (e.g. Biber, 1988; Biber 2004), texts have frequently been considered as related along various situational or functional parameters, such as interactive/non-interactive and narrative/non-narrative. These parameters are continuums of variation - texts can be more or less narrative.

With respect to the language of Twitter, Sardinha (2014) compared tweets and other Internet registers to the major patterns of linguistic variation found across pre-internet registers, as described in Biber's (1988) original MDA of spoken and written English. Sardinha (2014) found that tweets tended to be more associated with an involved communicative function, characterised by a high frequency of pronouns and reduced forms like contractions. Out of the Internet registers, tweets were most similar to emails and Facebook posts with respect to this pattern of linguistic variation, and they were least like blogs and webpages, as these were more informational. The next major pattern of linguistic variation of pre-internet registers relates to the degree of narrativity, and tweets were found to be more associated with the non-narrative side of the dimension. On the next dimension, tweets were marked for being the Internet register that was most associated with a situation dependent reference, characterised by a high level of place and time adverbs. With respect to this dimension, tweets were most similar to the pre-internet register of telephone conversations. The fourth dimension of linguistic variation refers to the overt expression of persuasion. Tweets were only slightly marked for persuasion. Finally, the next major dimension of linguistic variation of pre-internet registers refers to an abstract communicative style, characterised by a high frequency of conjuncts, passives and adverbial subordinators, which are associated with texts with high levels of technical content and complex logical relations. Tweets were slightly associated with an abstract style and were most like popular lore registers in this regard. Overall, Sardinha (2014) likened tweets to the digital equivalent of spoken language, as it was found that tweets clustered with other kinds of spoken texts, such as spontaneous speeches, face-to-face and telephone conversations. Whilst this research reveals where tweets fall in relation to patterns of linguistic variation of pre-internet registers (e.g. fiction and press reportage), other research has begun to describe the patterns of linguistic variation across internet registers more specifically.

Using Biber's (1988) approach, Titak and Roberson (2013) analysed a corpus of online registers, including Facebook and Twitter posts, emails, blogs and reader comments and revealed the major patterns of linguistic variation found across them and discussed the overall tendencies of each register. In their study, Facebook and Twitter posts were grouped together and were revealed to be associated with a descriptive informational production and were most similar to online newspaper articles along this parameter of language variation. Additionally, Facebook and Twitter posts also had a moderate level of involved and interactive discourse similar to blogs. Facebook and Twitter posts were not associated with the next dimension, which was interpreted as the complex statement of opinion. Finally, Facebook and Twitter posts were most strongly associated with a present tense orientation, as opposed to past tense. Overall, they found that Facebook posts and tweets, which can be characterised as microblogs, differ considerably from ordinary blogs, which are more narrative and more elaborated. They attribute this finding in part to the character restriction imposed on tweets.

Whilst Titak and Roberson (2013) grouped Facebook and Twitter posts together, Friginal, Waugh and Titak (2018) and Sardinha (2018) distinguished them. In particular, Sardinha (2018) also analysed the major patterns of linguistic variation across online registers, although in his study he included tweets, blogs, webpages, Facebook posts, and emails. Three major dimensions of linguistic variation were found. Tweets were found to be more associated with the positive side of Dimension 1, which is characterised by an involved and interactive function. Sardinha (2018) interprets the second dimension of linguistic variation as reflecting the “Expression of stance: Interactional evidentiality”, where the texts most associated with this dimension tend to be expressing their attitude towards knowledge. Tweets are moderately associated to this pattern of variation, although emails are most strongly associated with it. The final dimension is also interpreted as reflecting the expression of stance. The kind of stance found in this dimension of linguistic variation concerns expressions of personal attitudes, emotions and feelings. Facebook posts were found to be most strongly associated with this dimension followed by emails and then tweets, which were only slightly associated.

Friginal, Waugh and Titak (2018) not only differentiated Facebook posts from Tweets, but they also separated each group further according to particular topics, including (1) politics, (2) business, (3) entertainment, (4) personal, (5) sports, and (6) weather, and then computed how associated each topic-specific and platform-specific corpus was to the dimensions of online registers revealed in Titak and Roberson (2013). They found that tweets generally were associated with a descriptive, informational production. Additionally, they found that tweets were generally less associated with an interactive and involved function, as well as a complex statement of opinion communicative function. Moreover, tweets were generally more associated with a present tense orientation. Nevertheless, despite these general patterns, they importantly found that each group of topic-specific tweets varied in relation to the dimensions. Tweets that were associated with a descriptive, informational production were those on the topics of business, weather and politics, whereas tweets on the topics of entertainment, sports and personal were more associated with a personal narrative focus. Similarly, tweets that were interactive and involved tended to be those on the topic of sports, entertainment and personal topics, whereas the weather, politics, and business were less interactive. Tweets on the topic of entertainment and politics were associated with a complex statement of opinion, whereas weather was least associated with this particular communicative function. Finally, political tweets were most associated with a past tense orientation, whereas tweets on personal topics and business had a present tense orientation. Overall, this study enabled the comparison of topic-specific tweets in relation to the major patterns of linguistic variation across online registers. Importantly, this study reveals that tweets can vary from one to the next in relation to the patterns of linguistic variation of online registers.

This growing body of research is aimed at revealing the major patterns of linguistic variation across online registers and has enabled rich descriptions of the ways in which online registers like tweets and even particular topic-focused groups of tweets compare to other online registers along the major dimensions of linguistic variation found across the online registers. Nevertheless, little is known about the range of linguistic variation found across Twitter specifically. The research reviewed here details the major tendencies of groups of tweets in relation to other registers; however little is known about how tweets vary from one to the next and the major linguistic repertoires found across Twitter. Whilst we know that there are a variety of different kinds of tweets for a variety of different purposes, it is not yet known whether these differences are realised in actual linguistic distinctions.

The present section has reviewed literature that demonstrates that Twitter is used for a variety of different purposes from personal expression, interpersonal interaction and debate, as well as a source for news and real-time information, and as a platform to contribute information and news on a global and local scale. Despite this literature, there has not yet been an examination of the full range of the major communicative functions of tweets. In particular, we know relatively little about the various linguistic repertoires, properties and communicative purposes found specifically on Twitter. Importantly, Twitter plays host to communication and social connectivity, which can bind communities and individuals together (Asenas and Hubble, 2018), especially those that come from all walks of life and potentially have dissimilar views, norms and expectations. Whilst being exposed to alternative views is generally beneficial to individuals, as it can lead to a re-evaluation of one’s own views in light of other opinions (Chiluwa and Ifukor, 2015), this is Twitter and its effects in its best form and this is not often the case (Asenas and Hubble, 2018). Specifically, there is a darker side of Twitter and its technological features and affordances, one that fosters inauthenticity, hostility and impulsivity, and leads to a lack of respectful listening.

2.2.3. The Dark Side of Twitter

As mentioned, the technological affordances and features and architectural characteristics of Twitter can influence communicative practices and conventions. For example, the box where users post their Tweets contains "What's happening?", and this encourages tweets to not only concern a wide range of phenomena, but at the same time, it might lead to an increase in posts that concern events happening in the here and now, and the relatively close past, as opposed to reporting on what happened several years ago. Moreover, the non-reciprocity of following and the '@' mention has enabled users to keep up-to-date with their favourite celebrities and directly communicate with them, which may lead to feelings of closeness and intimacy. This view derives from media ecology (Meyrowitz, 1994), specifically the defining features of the communication technology influences how users of the medium process and make sense of information and how this influences behaviour. Additionally, social semiotics is a field focused on exploring the interrelation between the technology and the social practices. Whilst Schmidt (2014) suggests that the features of Twitter enable the emergence of personal publics, which he suggests provide opportunities for participation and social inclusion, other research (e.g Ott, 2017) has examined the platform Twitter for its defining features and argued that whilst much of what is posted on Twitter is innocuous, Twitter, nonetheless, fosters simplicity, impulsivity and incivility. Specifically, Ott (2017) suggests that tweets can be argumentatively simple due to the character restrictions, meaning that Twitter excludes complexity, and as a result tweets lack detail and sophistication. Ott (2017) suggests that such a characteristic undermines Twitter users' capacity to discuss and think about issues and events in complex ways. Whilst Ott's (2017) argument is in some cases true, it is important to note that individuals can create threads and tweet multiple times to extend their argument, thus the character restriction can be bypassed in some ways, meaning that tweets can be argumentatively complex; nevertheless, this may be rare.

Extending the view that tweets do not demand complexity, Ott (2017) suggests that tweets are consequently impulsive due to the lack of effort required in tweeting and the ability to post on-the-go through mobile devices, which means that the whole process lacks forethought, reflection, and consideration, especially of the consequences of the message. Rather, tweets end up being emotionally charged and eccentric. Recent research has found that impulsivity is a key factor in problematic media use (Orosz et al., 2016), where impulsive people, especially those who have been provoked, were more likely to comment immediately, often in a negative and uncivil manner (Koban et al., 2018). Arancibia and Montecino (2017) also describe how an event that produces frustration amongst Internet users instigates reactive and aggressive interaction on social media platforms, such as Twitter. Additionally, such impulsive and emotionally charged tweets have been found to receive more retweets in comparison to neutral tweets (Stieglitz and Dang-Xuan, 2013), which arguably means that aggressive tweets are ‘rewarded’ on Twitter, and are thus encouraged - a sentiment which has also been noted by Anderson (2019).

Finally, Ott (2017) argues that tweets are uncivil because they are informal, and because Twitter depersonalises interaction. Specifically, individuals do not consider how their tweets affect others, meaning that they will be more likely to post hurtful things than if they were forced to say the same thing to the individual’s face, a phenomenon known in psychology as ‘deindividuation’ (Siegel et al., 1986; Lea and Spears, 1991, 1992), which is also increased on Twitter because individuals have the ability to be anonymous (Hardaker, 2017; Hardaker and McGlashan, 2016), although research examining whether anonymity influences more hostile communication has been divided.

Wulczyn et al. (2016), for example, found that users were more likely to attack if they were anonymous, although non-anonymous users produced the majority of the attacks. Additionally, Omernick and Sood (2013) found that anonymous users produced more swearing, anger and negative emotion words than non-anonymous users and that non-anonymous users produced content that was more relevant to the topic being discussed than anonymous users. Alternatively, in a study investigating the effect of sanctioning racist commenters online, Munger (2016) found that more anonymous individuals were more likely to change their behaviour after sanctioning, although in a later study on uncivil commenters (Munger 2017), the opposite was found, suggesting that anonymity and the function of comments (uncivil or racist) are intertwined. Whilst anonymity may influence some cases of anti-social and malicious behaviour, it does not account for environments where individuals are anonymous and yet produce very civil discourse, nor does it account for environments where individuals are not anonymous and yet produce a substantial amount of abusive content (Miller, 2012). Whilst much of this research notes the negative effects of anonymity, the ability to be anonymous on Twitter has provided individuals with the tools to speak out about corruption and other negatively marked events, which would not be possible without anonymity or would have serious negative consequences for the individual (Arancibia and Montecino, 2017). Anonymity also allows individuals to explore their identity online and do positive things, as well as negative things (Hogan, 2013), although the degree to which something is positive or negative depends on who is observing.

In addition to uncivil language, other negatively marked linguistic acts have been found to be fostered on Twitter. For example, Nicotra (2016) suggests that Twitter facilitates public shaming, Dumenco (2011) shows that Twitter encourages spreading half-truths, misinformation and lies, especially because there is nothing demanding that the individual needs to tell the truth, and Demirhan and Çakir-Demirhan (2014) claim that patriarchal discourse on the social roles of women is perpetuated on Twitter, predominantly due to the fact that the alternative discourse is limited in number. Whilst these acts are hardly new, the communication technology makes it far easier and more accessible than ever before.

Another activity associated with the darker side of Twitter are firestorms or shitstorms, which refer to cases where an individual, group, institution or organisation suddenly receives a large amount of negative attention on Twitter, although research suggests firestorms have little long-term impact with respect to patterns of discussion post-firestorm (Lamba, Malik and Pfeffer, 2015). Nevertheless, the ability to create, contribute and spread negative remarks or reviews about an individual, company or product is made exceptionally easier on Twitter and such comments are not only confined to individuals within one's offline social network, but rather can reach their online social network and further, as most Tweets are posted publicly and therefore can be accessed by anyone with an internet connection.

Moreover, when many people start talking about the same topic, then Twitter makes it a trending topic, which makes it more accessible. Thus, whatever or whoever the subject is of a firestorm or shitstorm tends to become a trending topic and this can lead to more people contributing and commenting and may make things worse. For example, Justine Sacco was subjected to these negative consequences, which not only led to her being publicly shamed and the subject of various negative comments, but it also led to her losing her job. Thus, whilst Twitter brings people closer together, removing the communication barriers previously existed between people all over the world from all walks of life, it provides a catch-22 of increasing the accessibility and spread of more sinister and uncivil varieties of language.

Celebrities, politicians and people of public interest use Twitter. This allows supporters and fans to communicate with them and vice versa, thereby removing communication barriers that have previously existed. Nevertheless, this has also been exploited for negative means. For example, Ouvrein et al. (2018) examined celebrity bashing in newspaper comments. Although not directly relevant to Twitter, celebrity bashing is also facilitated on Twitter. This is because people frequently share newspaper stories via URLs on Twitter, often adding commentary. Whilst this commentary can be innocuous, it is also very easy for it to be insulting to the celebrity discussed. Previous cases, such as Lesley Jones who was targeted on Twitter by racist trolls following the Ghostbusters' film premiere (Woolf, 2016) shows how the ability to communicate with celebrities on Twitter or interact with their tweets and news stories about them can facilitate celebrity bashing.

Individuals with diverse viewpoints and opinions use Twitter. Twitter can thus be instrumental for observing and engaging with alternative viewpoints. Nevertheless, Dahlberg (2001) disputes that the Internet provides a space for rational-critical discourse, and this can arguably be applied to Twitter especially considering some of the research above. Specifically, Dahlberg (2001) examined the extent to which the Internet provides a space for rational-critical discourse by using Habermas' theory of democratic communication. While Dahlberg (2001) demonstrates that political claims are exchanged and critiqued, the quality of such discourse does not entirely match the public sphere model. With respect to the current research, Dahlberg (2001) demonstrates that there is a lack of respectful listening to others, identity claims and information put forward is not always verifiable, and the social inequalities in the offline world are reaffirmed online with certain individuals and groups dominating much of the discourse. Similar patterns have been observed in the past on Twitter and social media. Specifically, fake news and the spread of mis- and dis-information became exceptionally severe throughout the campaign for Britain to leave the European Union, and throughout the 2016 U.S. presidential election. Additionally, throughout these political debates, anti-liberal and racist groups like the Alt-right gained importance and entered mainstream knowledge.

These studies suggest that the architecture and conventions of Twitter may cultivate tweets to possess certain characteristics, which can lead to a variety of negative communicative practices, and which can be used for negative means. Overall, this section reveals that research suggests that most people use Twitter to post innocuous feelings, opinions, ideas and information to their followers, however there is a growing number of studies showing that there are some individuals who exploit the platform to promote a hostile environment. These individuals and their behaviour have been labelled in a variety of ways, but most commonly, they are called *trolls*, and their behaviour is deemed *trolling*.

2.3. Trolling

2.3.1. The Diversity of Trolling

> *!!Can somebody explain to me what "trolling" is!? Because .. *whistles* I'm a little lost. I hear it used different ways every time.*
>
> (@youNEAR_UHgee, 2016)

Trolling initially referred to an experienced user in a newsgroup posting an erroneous or exaggerated message on a previously discussed topic in order to provoke a new member to post a follow-up article that points out the misconception (NetLingo, 1995-2015). Despite this original meaning, and as the tweet above and the following examples illustrate, *trolling has* since been used in various ways. For example, European Commission spokesman, Margaritis Schinas was recently described on Twitter as an expert in Brexit trolling (@Cromwell606, 2019), when he creatively and unusually quoted the Spice Girls' 'Wannabe' song in a frustrated and exacerbated plea to the UK to "tell us what they want, what they really, really want" in the Brexit negotiations.

Alternatively, the media referred to a group of trolls who have recently targeted Khadija Ben Hamou with a deluge of racist abuse online after she became the first black woman to win the title of Miss Algeria (Prideaux, 2019). Additionally, anonymous cyber-bullies and trolls targeted Brandy Vela for being overweight (Salo, 2016). The abuse led to Brandy eventually committing suicide by shooting herself in front of her parents. Despite her death, the abuse continued on a Facebook memorial page for Brandy, where trolls posted malicious abuse and included pictures of a pig with Brandy's face on it, and also a picture of Brandy with guns (Salo, 2016).

In another case, a tweet in November 2018 was posted from @Complex calling for screenshots of an interaction where children asked their mothers "how long does it take to microwave a turkey?" for thanksgiving dinner. The responses ranged from the children's mother's innocently assuming they were defrosting the turkey, suggesting approximate defrosting times, to expressions of shock and clarification that "you do not microwave turkeys, you bake them", and even an irrelevant but yet catastrophic response of "I'm leaving your father". The media labelled this prank as *trolling* (Anderson, 2018).

These wide-ranging uses of *trolling* in the media and online illustrate that, depending on the context, *trolling* can mean anything from hate speech and malicious cyber-bullying to pranks and playful banter. One reason for this may be because one aim of trolling is to provoke a reaction, and being provoked is a personal reaction. Thus, trolls with this communicative goal may adapt their behaviour according to what is most likely to provoke a response from their target. Because of this motive and other reasons, *trolling has* developed into a multi-faceted phenomenon that for the most part is operating as an all-encompassing term for most negative (Hardaker, 2010; 2017) and/or socially transgressive (Phillips and Milner, 2017) behaviour online (and occasionally offline), in the sense that the (online) community or platform negatively marks the behaviours because they go against what is perceived to be ‘normal’ behaviour.

Norms are group-specific habitual behaviours. The norms of online platforms and communities within platforms are sometimes detailed explicitly (Pavalanathan et al., 2018). Reddit, for example, has its own detailed guidelines and etiquette, termed ‘reddiquette’ (https://www.reddit.com/wiki/reddiquette), as well as guidelines for each individual subreddit (Fiesler et al., 2018). Other platforms and communities, however, do not set out the norms so explicitly. For example, Twitter does not have any formal guidelines or collection of rules indicating etiquette because Twitter is largely made up of millions of communities. For some communities on Twitter anything goes, although this is not always the case in other communities on Twitter, where individuals who violate group norms will be subject to suggestions or criticisms and potentially ostracism (i.e. they may lose followers or be banned following other Twitter users reporting them and from Twitter moderation). These suggestions and criticisms are where the rules or etiquette of tweeting for that particular community emerge and are constructed (Schmidt, 2014; Pavalanathan et al., 2018). Consequently, in communities where there is no formally codified set of rules, the norms of a community are constructed socially, emerging through interaction of the members (Burnett and Bonnici, 2003), as well as through different types of moderation and regulation, often as a result of disruptive behaviours (McLaughlin et al., 1995; Herring, 2002).

Because trolling largely violates what is perceived to be ‘normal’ behaviour, the behaviours of trolling are often platform- and/or community-specific because it is the rules and norms of a platform or community that influence which behaviours are considered deviant. For example, trolls have been constructed as a threat in some areas of academic research (Lundberg et al. 2018). Specifically, Lundberg et al. (2018) acknowledge the potential that trolls could attempt to re-identify data that has been anonymised in datasets that have been released in an attempt to attack the survey respondents, as well as attack the researcher’s integrity. This view that it is deviant to re-identify anonymised data is not held by all communities. For example, determining who the likely author is of an anonymous disputed text is not considered to be deviant in the forensic linguistic community in a criminal investigation demanding an authorship analysis, such as in the murder investigation of Amanda Birks (see Grant, 2013). Thus, it is argued that trolls display variation in the behaviours and strategies according to community-specific norms and values, which they firstly learn and assimilate towards, and then purposefully transgress (Cruz et al., 2018).

As a result, studies have chosen to investigate trolling by focusing on its occurrence within particular social media sites, web-based communities and/or social networks. For example, trolling and other non-normative behaviours have been examined on Usenet (Hardaker, 2010; 2013; 2015), 4chan (Manivannan, 2013; Milner, 2013; Higgin, 2013; Hine et al., 2017), Reddit (Merritt, 2012; Mojica, 2017), YouTube (McCosker, 2013; McCosker, 2014), Twitter (Hardaker and McGlashan, 2016; Synnott et al., 2017; Abril, 2018), Facebook (Phillips, 2016; Karppi, 2013; Ditrich and Sassenberg, 2017; Buglass et al., 2016), Newspaper website comments (Jones, 2013), Wikipedia comments (Shachaf and Hara, 2010), virtual games, such as 'World Of Warcraft' (Higgin, 2013), video chat environments, such as 'Chatroulette' (Kopecký, 2016), as well as Location Based Real Time Dating (LBRTD) apps, such as Tinder® (e.g. March et al. 2017). These studies show that trolling is ubiquitous and far-reaching and that many of the behaviours are at odds with community norms. For example, Kopecký (2016) describes how trolls go against normative behaviours in video-chat environments like Skype or Chatroulette and misuse webcams. Specifically, webcam trolls use a forged video loop and introduce themselves under a fake identity to children to convince them to expose themselves on webcam, which is subsequently recorded and then used later for a variety of means, such as for blackmail purposes.

One way to examine trolling in particular web-based communities can be to observe how individuals within the web-based community talk about and perceive trolling. For example, Hardaker (2010) analysed and coded over 2,000 UseNet posts containing the word 'troll' or some variation (e.g. *trolling*) from particular discussion lists for the users main point or issue, revealing that trolls were largely discussed on these lists with respect to their *deception*, *aggression*, *disruption* and *success*. Specifically, Hardaker (2010) found support for Donath (1999) and Dahlberg's (2001) description of trolling, which is that trolls will try to deceive the community into thinking they are a legitimate community member, and then they will attempt to disrupt the community, whilst trying to keep up the facade of appearing as a genuine member. Hardaker (2010) found that other users discussed trolls according to their provocative nature because they were either aggressive by insulting or attacking others, or because they disrupted the flow of conversation by posting irrelevant or repetitive posts. Finally, Hardaker (2010) revealed that users discussed the troller’s success and constructed what was a successful and unsuccessful troll.

Other research has also examined perceptions of trolling, toxic behaviour and cyber-bullying (e.g. Sanfilippo, Yang and Fichman, 2017; Kwak et al., 2015; Marwick and boyd 2011). The results show that perceptions vary depending on whether one is a victim or observer, as well as what platform the trolling posts appear on, and also what age the perceivers are. For example, Kwak et al. (2015) found that toxic behaviour and what constitutes it was perceived differently by victims and third-party bystanders, as well as between cultures. Additionally, previous research demonstrates that young people view bullying differently from adults (Marwick and boyd 2011).

In a study designed to understand how college students perceive trolling, Sanfilippo, Yang and Fichman (2017) found that their participants argued that some trolling behaviours on particular platforms are not problematic and that they should not be treated equally because they perceive trolling to be a diverse phenomenon. In particular, they found that their participants perceived some ideological trolling to be necessary. Maltby et al. (2015) also examined students’ perceptions of trolling behaviours and found that online trolls are perceived to have low self-confidence, and to be attention-seeking, vicious, funny and uneducated. By contrast, Cruz et al.

(2018) found that their participants perceived trolls to be educated, especially educated and well-informed on the community and community-specific topics, because they felt that an uneducated troll's post would be idiotic and not have the same effect as it does when they have knowledge about the particular community.

Similar to Hardaker (2010) who examined the discursive construction of trolling and trolling strategies, Petykó (2017) also explored what motives are attributed to trolls in comments on three political blogs. Petykó (2017) found five motives. One of the motives for trolling that the commenters discussed was because of a broad spectrum of emotional states, mental health issues and social problems. Another motive for trolling that the commenters described was that they were being paid to troll, although it was not specified who was paying them. The third and fourth motive described by commenters was because of their affiliation to particular political parties or because of their support for a political party or ideology (e.g. *Tory troll*). The final motive specified was that trolls were working for either a political body, country or under the instruction of the European Union.

In addition to interviews and responses to trolling, perceptions of trolling have also been analysed in threads discussing trolling. Coles and West (2016) analysed responses to an academic report (Buckels et al., 2014 see below) investigating trolling. Specifically, the discussion involved individuals disputing, discussing and negotiating what trolling and trolls are. Coles and West (2016) found four repertoires of discussions of trolling, which noted that (1) trolls are easy to identify; (2) trolls are harmless and different nowadays to how they used to be; (3) trolls need to be identified and responded to appropriately through trolling the trolls, and (4) trolls are nasty and being a troll is not acceptable. Whilst this research reveals how trolling is defined by third-party bystanders and victims, other research has explored how trolls define themselves.

In interviews with trolls in gaming, Cook et al. (2018) asked them how they define trolling and what sort of behaviours constitute it. Whilst there was considerable variation, they found that the trolls stressed three consistent non-mutually exclusive elements. These were: (1) attack - directly attacking other people's enjoyment of the game; (2) sensation-seeking - the enjoyment of obtaining a reaction from the victim, and; (3) interaction-seeking - the desire to interact with other participants for friendship. Although informative for trolls in gaming, it is not yet clear how generalisable these findings are to trolls on other platforms and online communities.

Overall, this collection of research shows that the perceptions of trolling and trolls vary along multiple dimensions, including its motives, its purposes, its strategies, its intelligence, its appearance, its effect, and how it should be responded to and managed, and this is dependent not only on the community or platform, but also based on who is doing the perceiving/defining (i.e. the trolls, victims, third-party bystanders, the media, young people and old people).

In addition to the various uses of *trolling in* the media and online in different communities, previous academic research shows that there is little agreement on what trolling is. Trolls have been defined as deceptive (Donath, 1999; de Seta, 2013) and hostile (Hardaker, 2010; Weinstein et al., 2015). Other definitions describe the posts of trolls as benign contrarian (Gaus, 2012), inflammatory and abusive (Nicol, 2012), malicious (Coles and West, 2016), provocative (Hardaker 2010; 2013; McCosker, 2014) or of a teasing nature (Mihaylov and Nakov, 2016). The literature has also focused on the purpose of trolling, for example, to expose new members (Hardaker, 2010), for amusement (Hardaker, 2015), 'for the lulz' (laughing at people's expense) (Phillips, 2011; 2013; 2016; Asenas and Hubble, 2018), for attention-seeking purposes (Hardaker, 2010; Shachaf and Hara, 2010; Cruz et al., 2018), to insult (Ansong et al., 2013), to disrupt (Ansong et al., 2013) and hijack the conversation (Lumsden and Morgan 2017), to obtain a reaction (McIntosh and Pavlik, 2011; MacKinnon and Zuckerman, 2012; Lumsden and Morgan, 2012), to antagonise (Klempka and Stimson, 2014), to distract and mislead (Hogan, 2013), to 'out' the victim from participation in public forums of debate (Lumsden and Morgan, 2017), and to manipulate opinions (Mihaylov et al., 2015), including public opinion (Nyst and Monaco, 2018; Fokin, 2016). Whilst several descriptions of trolling have been offered, there is yet to be a systematic analysis of the extent of its linguistic repertoires, properties and communicative functions.

Because of its widespread use as a behavioural catch-all, some researchers have attempted to differentiate trolling from other negative online behaviours such as flaming (Hardaker, 2017; Hmielowski, Hutchens and Cicchirillo, 2014; Abril, 2018; Asenas and Hubble, 2018), cyber-bullying (Vandebosch and van Cleemput, 2008; Golf-Papez and Veer, 2017; Guy and Shapira, 2018), and hate-speech (Phillips and Milner, 2017; Zhang, Robinson and Tepper, 2018). For example, Hmielowski, Hutchens and Cicchirillo (2014) suggest that the purpose of trolling is to waste the time of their victims, but this needs to be discrete, which is unlike flaming, where the victims know the purpose of the abuse. Flaming has been described as a much more aggressive kind of abuse than trolling (Bacile et al., 2018). Specifically, flaming comprises a high level of insults, swearing, hate and hostility (Bacile et al., 2018). Alternatively, Abril (2018) suggests that successful trolling requires a response; otherwise it is flaming or unbitten bait. This suggests that Abril (2018) perceives that the purpose of trolling is to retrieve a response, whereas flaming does not share the same purpose. Cook et al. (2018) also make this distinction that trolling and flaming differ in their purposes, although flaming is able to cross over into trolling when the aim is to obtain a response. Additionally, Asenas and Hubble (2018) suggest that trolls are motivated by ‘lulz’, which is an aggressive form of laughing at other people’s expense, whereas the motivations behind flaming are not the same, although they do not elucidate what the motivations of flaming are.

In addition to flaming, trolling is also distinguished from cyber-bullying. Specifically, the attacks constitute cyber-bullying when they are repetitive and are to individuals that are younger (Vandebosch and van Cleemput, 2008; Golf-Papez and Veer, 2017; Guy and Shapira, 2018), suggesting that trolling is not repetitive and that age differences are irrelevant for trolling. Zhang, Robinson and Tepper (2018) suggest that the purpose of cyber-bullying is to harass, threaten or intimidate individuals rather than groups and this compares to hate speech, which tends to target individuals or groups on the basis of their characteristics with the aim of inciting harm or promoting hatred.

Hate speech is also distinguished from trolling by Phillips and Milner (2017) for ethical reasons. Specifically, they emphasise that there will be negative consequences of grouping them together or by labelling instances of hate speech as trolling, such as the desensitisation of hate speech. Whilst these distinctions are made for a variety of purposes, they are often dependent on the researchers aims (or ethics). However, there lacks agreement and there are often multiple grey areas (Cruz et al., 2018). Fundamentally, many of these distinctions run counter to everyday uses of the terms. For example, it has been revealed that individuals online are using *trolling as* an umbrella term to refer to a variety of negatively marked behaviours (Hardaker, 2010; Hardaker, 2013; Hardaker, 2015; Hardaker and McGlashan, 2016), including flaming, cyber-bullying (Miller, 2012; Lumsden and Morgan, 2017), and hate speech (Clarke, 2019). Thus, many have sought to specify the particular behaviours it is used to label, and by extension, the different types of trolling.

For most researchers, there is *behavioural*, *visual* and *verbal* trolling (Cook, Schaafsmer and Anteunis, 2018; Veszelszki, 2017). *Behavioural trolling* largely refers to that conducted within gaming, such as killing teammates or doing something that compromises the team's success, like walking away from the keyboard, allowing oneself to be killed by the opposing team (Cook, Schaafsmer and Anteunis, 2018), and even the more extreme practice of swatting, which refers to cases where individuals find out another player's address and anonymously call the police informing them of a crime at the address, leading to a Special Weapons and Tactics (SWAT) team being sent there, essentially stopping the individual from playing the game.

There is also *visual trolling* and examples of this include posting memes, or sending unrequested explicit photos of oneself to another, deemed the digitised version of ‘flashing’ or ‘naked’ trolling (Lumsden and Morgan, 2017), as well as posting photos that misrepresent the truth, such as when an individual posted a picture of a piece of fried chicken claiming that it was a deep fried rat because it unfortunately looked like one (Veszelszki, 2017). Mocanu et al. (2014) also found that trolls posted controversial, parodistic and satirical content, often memetic, which was aimed at mimicking the way that alternative news sources fabricate highly fictitious statements. This mimicry nature has also been described by Phillips (2013; 2016) but towards mainstream media, as opposed to alternative media (discussed below).

Misrepresentation and fabrication does not only occur visually, but also verbally. *Verbal trolling* refers to the linguistic act, and this can encapsulate a variety of communicative acts, including uncivil acts, such as mockery, insults (Wang and Silva, 2018), vitriolic abuse (Synnott et al., 2017), and rape and death threats (Lumsden and Morgan, 2017), as well as posting nonsensical content (Synnott et al., 2017), vandalising Wikipedia pages (i.e. changing the content on pages) (Alkharashi and Jose, 2018), and/or spreading fake news, misinformation and disinformation, and this can be for the purpose of manipulating opinions (Nyst and Monaco, 2018; Fokin, 2016). For example, state-sponsored or ‘hybrid’ trolls are instructed by a particular state or state institution to communicate a particular ideology or target individuals critical of the state with hate and harassment in order to intimidate and silence them (Nyst and Monaco, 2018; Fokin, 2016). Such state-sponsored trolls are also instructed to spread disinformation, which is not necessarily aimed at persuading people, although this can happen, but rather it is aimed at complicating the context and knowledge by polluting it with an abundance of competing, false and useless information (Fokin, 2016), which can ultimately instil doubt in common knowledge and drive conspiracy. Hybrid or state-sponsored trolling are distinguished from classic trolling in that the latter trolls for personal motivations, whereas the former is acting on behalf of other individuals, groups or the state (Fokin, 2016), although determining whose interests the trolls are acting on is impossible. Another common type of verbal trolling found in gaming was referred to as trash-talking, which is comparable to one-upmanship in sporting contexts (e.g. *we are going to win, we are the best team*), although there were more negative cases, including direct criticisms, insults or degrading the player's family members (e.g. *your mum is a cunt*, *you are stupid*) (Cook et al., 2018).

In addition to functioning as an umbrella term, encapsulating a whole range of behaviours, communicative acts and types, trolling has been positioned as a type of broader behaviour. For example, Phillips (2013; 2016) argues that trolling is reflective of mainstream media practices. Specifically, Phillips (2016) notes that trolls exploit and hijack the media's practices of spectacle, sensationalism and hyperbole for success and profit in a process of détournement, whereby the meaning of a particular entity is turned against itself. Phillips (2016: 68) finds that trolls détourn "corporate media strategies for explicitly lulzy ends" in order to "reinforce the real meaning of an original element" (Jappe, 1999: 59). "Trolls troll Fox News by acting like Fox News[…] then howl with laughter when their chosen targets unwittingly rail against their own reflections" (Phillips, 2013: 506). In essence, the media's behaviours tend to be accepted as being part and parcel for profit-making, and yet trolling behaviours, which are exactly the same, are condemned and negatively marked (Phillips, 2016).

Additionally, trolling has been positioned as a type of griefing in gaming (Rubin and Camm, 2011), which refers to the situation when a gamer does not complete the tasks of the game, but instead disrupts the enjoyment of their opponents game and intends to cause grief (Coyne et al., 2009; Dibbell, 2009; Foo and Koivisto, 2004). Trolling has also been positioned as a type of online incivility (Bacile et al., 2018), although these categorisations fail to acknowledge the more playful and harmless side of trolling that can be experienced on both sides, like Phillips' (2016). There are also competing labels for similar types of behaviours. For example, individuals who violate norms through disagreeable and/or unsociable behaviour in online social networks have been referred to as *trolls* (Phillips, 2016) and *troublemakers* (Buglass et al., 2016). Whilst this backdrop of definitions, labels and descriptions might appear to complicate things, it importantly reflects that trolling is not a one-dimensional phenomenon, but rather it is a multifaceted and often subjective construct – one that is continuously evolving and adapting.

2.3.2. Why Do People Troll?

In addition to defining what *trolling* is and its purposes, strategies and types, other research has aimed to understand *why people* troll, which many have ascribed to the ability to be anonymous online after the theory of deindividuation. Essentially, deindividuation theory suggests that when individuals are immersed in a crowd or group, they lose a sense of their personal identity (Festinger et al., 1952; Zimbardo, 1969; Diener, 1980). This loss of self-awareness means that individuals are more inclined to act aggressively and anti-normatively (Festinger et al., 1952; Zimbardo, 1969; Diener, 1980). Zimbardo (1969) emphasised that anonymity was another cause for loss of self-awareness. As a result, trolling has been said to occur because of the anonymity, which the internet affords (Hardaker, 2010; Herring et al., 2002).

Nevertheless, and as mentioned in section [2.2.3], studies examining this assumption have found differing results. Importantly, this theory does not account for individuals who are not anonymous yet still act aggressively and anti-normatively online. To account for these exceptions, Miller (2012) argues that trolling and other forms of antisocial and harmful behaviours online are fundamental examples of the effect of our understanding that care and responsibility to others are based on physical proximity to them, as opposed to a mediated closeness. In other words, despite increased connectivity, which results in a sense that spatial limitations are less problematic (i.e. we can communicate with anyone in the world with an internet connection), the essence of an ethical social encounter is, according to Miller (2012), dependent on bodily face-to-face interaction and not mediated, disembodied interaction. Thus, Miller is arguing that individuals are more inclined to exhibit antisocial behaviour online because their interlocutor is not physically present and this is because the technology affords mediated, disembodied interaction, as opposed to face-to-face interaction. Although convincing, there are still exceptions to this; that is, when individuals are abusive when they are face-to-face with someone.

Other explanations for why people troll have been provided in the literature and these tend to fall into the nature versus nurture debate. For example, some research has found that trolling and other anti-social behaviours are a result of the environment, including as a result of a negative context, negative mood, or the occurrence of other negative comments (Cheng et al., 2017; Wulczyn et al., 2016). Other research, however, has described that the motivations (Baker, 2001; Herring et al., 2002; Shachaf and Hara, 2010; Craker and March, 2016) and personality traits (Buckels et al., 2014; March et al., 2017; Sest and March, 2017; Barnes et al., 2018; March, 2019) of people who troll are distinctive to them.

For example, Wulczyn et al. (2016) examined the environment of personal attacks on Wikipedia comments by specifically looking at the number of comments that came before and after a personal attack, as well as non-attacking comments to evaluate whether personal attacks cluster in time. They found a significant difference between non-attacking comments and attacking comments, with attacking comments being more contextually close to other attacking comments than non-attacking comments. This was even prevalent in a small amount of neighbouring comments, indicating that personal attacks cluster in time in this domain. This suggests that users of sites might be more likely to post attacking comments following attacking comments.

Cheng et al. (2017) also found a similar result in their research. Specifically, this study found that more controversial topics induced trolling, and if the discussion also included trolling posts beforehand, then the likelihood of future troll posts increased. Additionally, trolling was more common late at night and early on in the week. Overall, they found that people were more inclined to exhibit antisocial behaviour if they were exposed to a negative context and were in a negative mood, which they suggest means that anyone can become a troll. These findings suggest that trolling is a spreadable phenomenon. Nevertheless, in an experimental study examining the impact of exposure to incivility on newspaper comments, Rosner, Winter and Kramer (2016) found little support for individuals posting uncivil comments after being exposed to increasing numbers of uncivil comments, although their realisations of uncivil and civil comments were arguably not different enough. Specifically, they manipulated messages to become uncivil often by including swear words, as opposed to understanding what lexico-grammatical features comprise a civil and uncivil style and then implementing these into method design.

Another study found that exposure to incivility increased aggressive cognitions, although it was not found to influence aggressive reactions such as in the form of posting back aggressively (Rosner, Winter and Kramer, 2016). Nevertheless, in their study on commenting on newspaper stories, Barnes et al. (2018) found that disagreeable individuals were more likely to comment on stories when they disagreed with the journalist and this may be trollish (Barnes et al., 2018).

Whilst this research has focused on the environment (i.e. what contexts are more conducive to internet trolling and antisocial behaviour), other research has sought to understand what psychological and motivational traits predict internet trolling behaviours. For example, Fox and Tang (2013) noted that in video game environments women received a large amount of misogynistic and sexist abuse. They investigated what personality variables predict sexist beliefs in these video game players. They found that social dominance, the desire for power over women and the need for heterosexual self-presentation were predictors of video game sexism. Overall, they found that participants who supported masculine norms were more likely to be sexist about women participating in video games, thereby supporting the expectation states theory, which suggests that individuals, specifically women, who act counter to the stereotypical, normative behaviours of their sex are likely to be penalised. This is apparent in gaming, where the default gamer is constructed as the white male (Gray, 2012a;b), and thus the presence of a female in video games is likely to be perceived as non-normative, and therefore they may be more likely to receive abuse in these environments. This was evident in Gamergate, which was a trolling movement of targeted harassment and threats to women in the video gaming industry (see Mortensen, 2018). For example, Brianna Wu received numerous threats on her and her family's lives because of her role within the tech industry.

In another study, Buckels et al. (2014) sought to examine the personality profiles of trolls. Specifically, they examined the personality traits known as the Dark Tetrad, which refers to the four personality types: narcissism, Machiavellianism, psychopathy, and sadism. They found that the Dark Tetrad was specific to trolling, especially sadism, where sadists tend to enjoy trolling. They also found that the enjoyment of other activities was not related to these personality traits (e.g. chatting and debating online).

Additionally, they found the frequency of posting, trolling enjoyment, and trolling identity and behaviour were strongly positively related, which they suggest supports the notion that excessive levels of social media use are linked to antisociality (Carr, 2011), although they note that the causal direction of these associations are not clear.

Craker and March (2016) extended this research to explore what motivational factors cause individuals to engage in trolling behaviour. Based on the understanding that the outcome is what determines engagement in certain behaviours, especially outcomes that are personally rewarding, Craker and March (2016) investigated the relationship between the Dark Tetrad personality traits, social rewards, specifically negative social potency, and Facebook trolling behaviours. Overall, they found that psychopathy and sadism were significant positive predictors of Facebook trolling behaviours, whilst Machiavellianism and narcissism were not. Craker and March (2016) also found that negative social potency was found to predict Facebook trolling behaviours, which they suggest means that individuals who partake in Facebook trolling behaviours are likely to be intrinsically motivated by obtaining negative power and influence over people as the social reward. Moreover, they found that sadism and psychopathy lost their significance in predicting Facebook trolling behaviours when negative social potency was added to the mix, which they suggest is because negative social potency is potentially the underlying reason for high levels of sadism and psychopathy. Thus, overall they argue that social motivation best predicts online trolling, as opposed to personality.

March et al. (2017) examined trolling on location-based real-time dating (LBRTD) apps. Building on the findings of Craker and March (2016), they found that LBRTD trolls demonstrated traits of psychopathy and sadism. However, they also examined impulsivity, specifically dysfunctional impulsivity as a potential predictor of trolling, which refers to the tendency to act without forethought. They found support for this, but also they found that moderate and high levels of psychopathy influence dysfunctional impulsivity.

Alternatively, Sest and March (2017) tested whether the association between psychopathy and trolling, found consistently in the studies above, could be attributed to a lack of empathy. They revealed that affective empathy, which refers to the ability to experience, internalise and respond to other people's emotions, was a significant negative predictor of trolling. They also found that psychopathy moderated cognitive empathy, which refers to the ability to recognise and understand the emotions of other people. Moreover, it was found that cognitive empathy was a significant positive predictor of trolling, only when the trolls displayed average to high levels of psychopathy.

Based on all of these findings, March (2019) aimed to combine all the individual predictors of trolling in one study, exploring the utility of all the psychological traits and motivations in previous research and additional ones to obtain a clearer understanding of trolling. Specifically, through questionnaires, March (2019) examined whether gender, primary psychopathy, sadism (direct and vicarious), affective empathy, cognitive empathy, negative social potency, and Vulnerable Dark Triad traits (i.e., secondary psychopathy, vulnerable narcissism, and borderline personality traits) could predict internet trolling. The results found no support that gender predicts trolling, which March (2019) suggests is because gender might be a context dependent variable, where trolling may be predicted by gender across some internet platforms and situations, and others it may not. March (2019) found support for Buckels et al. (2014), Craker and March (2016) and Sest and March (2017), which found that internet trolling is predicted by primary psychopathic and sadistic traits. Importantly, this study examined the effect of primary and secondary psychopathy and direct and vicarious sadism. Whilst internet trolling was predicted by primary psychopathy, March (2019) found that secondary psychopathy was not a significant predictor of internet trolling, which suggests that trolls are less impulsive, neurotic and emotionally reactive, but are instead more callous, manipulative and generally lack remorse (Levenson, Kiehl and Fitzpatrick, 1995).

March (2019) also found that both direct and vicarious sadism were significant positive predictors of trolling, which suggests that trolls not only enjoy hurting and humiliating others but also enjoy watching others get hurt or humiliated. Additionally, March's (2019) results support the findings of Craker and March (2016) that internet trolls are motivated by negative social potency. Similar to Sest and March (2017), it was found that affective empathy was negatively associated to internet trolling and that cognitive empathy and trolling were associated only when psychopathy scores were high. These results suggest that trolls are able to predict what will cause others to be distressed, whilst remaining detached from the emotional experience. Phillips (2011) similarly suggests that trolls are more likely to be detached and will avoid expressing how they feel because trolls ridicule such characteristics in others.

Overall, the results found no support that the Vulnerable Dark Triad predicts trolling, which means that the troll is not necessarily insecure, but rather that their self-worth is not contingent on the recognition of others – trolls are merely behaving the way they do because they enjoy it. This enjoyment has been supported in other studies. For example, Cook et al. (2018) investigated trolling goals and found that trolls said that they trolled because of personal enjoyment, for revenge, and to seek thrills. More recently, in their study examining the intentions, motivations and traits behind more subtle forms of cyber-aggression, Koban et al. (2018) found no relation to the Dark Triad traits. However, they found that individuals who became bored more quickly were more likely to consider responding to an uncivil comment with incivility. This suggests that the individuals are being uncivil for entertainment (i.e. to stop being bored). Additionally, Koban et al. (2018) found that excessive users of Facebook tend to consider producing more uncivil comments in controversial discussions, which they suggest could be due to familiarity with the conventions or desensitisation to such toxicity.

Importantly, these discrepancies within this band of research examining the environment, motivations, and personality variables of trolls and trolling comments can be explained by the understanding that trolling is a multi-faceted and variable phenomenon. Because trolling can vary so much, especially according to the community and/or context in which it resides, it is not surprising that different types of trolls and trolling behaviours on various platforms might be influenced by contexts, environments, motivations or personality traits, which are at variance with other trolls and trolling behaviours.

2.3.3. The Effect of Trolling

Cases like Brianna Wu described above illustrate that trolling has the potential for many negative outcomes. These negative effects can include interrupting the flow of a conversation and preventing future discussions (Golder and Donath, 2004; Abril, 2018), as well as driving people off particular spaces of the internet, such as Sara Payne (mother of a murdered school girl, who was harassed on Twitter by trolls), and Zelda Williams (Robin Williams' daughter who was also abused following the death of her father) (Cohen, 2014). Additionally, trolling has caused people to damage their own property (see Golf-Papez and Veer, 2017), such as when individuals were told that deleting 'System32' off their computer would speed up processing time, when in actual fact this makes the computer turn into a brick (Phillips and Milner, 2017), or when individuals were tricked into microwaving their phone after a fake advert was created which suggested that you could charge it in the microwave (Radulova, 2014). Trolls have also caused people to fear for their lives, such as Luciana Berger (a Labour MP), who was sent threatening and anti-Semitic messages from a previously convicted Twitter troll, such as "you will get it like Jo Cox" (another Labour MP, who was murdered by a Neo-Nazi) (Laville, 2017). There are also some types of trolling that have the potential for, and have had fatal consequences. Charlotte Dawson (see Morrissey and Yell, 2016), Brandy Vela and Amanda Todd are just three examples of individuals who have killed themselves as a result of being victims of trolling, cyber-bullying and online harassment. Additionally, Andrew Finch was the first fatality of swatting, where a police officer of a swat team killed Finch after being sent to his address following a hoax 911 call, which was later revealed to be from Tyler Bariss (Hellmore, 2018).

Despite such reports showing the harmful and sometimes catastrophic outcomes of trolling and its different guises, it also has the potential for being benign and instrumental. For example, Phillips (2016) shows that people are driven to sites where trolling is occurring so that they can watch it happen and/or be a part of it. As a result, trolling becomes beneficial and instrumental for such site owners, who gain revenue from advertisers who pay exceptional amounts to have their adverts on these particular sites. Additionally, the companies whose adverts are displayed might also benefit from trolling through increased awareness of products, and by extension an increase in sales. Moreover, previous research has demonstrated how some trolls troll in order to seek friendships and interact with people, often other trolls (Cook et al., 2018). For example, it has been shown that some trolls act together in organised attacks, such as Operation Google – trolls came together to replace smears for black people, Jews and Muslims with Google and Google related products, so that it would force the AI program to censor Google's name (Hine et al., 2017) – and in raids in the virtual world like World of Warcraft – trolls created black avatars and drew on African American stereotypes and came together to disrupt a community (Higgin, 2013). These examples show how trolls form communities and friendships around trolling successes and trollish activity, whilst simultaneously being antagonistic and exceptionally racist and abusive.

It is these sorts of divergent outcomes, responses and perceptions that led Phillips and Milner (2017) to suggest that trolling is *ambivalent* –

concurrently social and antisocial, weird and normal, desirable and undesirable, group-including and group-excluding, funny and insulting, creative and disruptive, clever and stupid – depending on who is participating and observing, when and where it occurred and what the particular intent of the poster was and what assumptions they brought to the interaction. The ambivalence of trolling is also arguably fuelled by the fact that the judgment of what constitutes ‘inappropriate’, ‘anti-normative’, ‘offensive’, ‘humorous’, ‘weird’, ‘antagonistic’, ‘social’, and ‘desirable’ is inherently a subjective one, partially influenced by community norms. Phillips and Milner (2017) describe how examples of trolling are too variable across specific cases for them to be essentialised as *this* rather than *that.* They also emphasise that trolling cannot be pinned to one purpose, but rather that each case inhabits the whole spectrum of purposes. Although each trolling instance can simultaneously be positive, negative and somewhere in between, the negative and devastating effects have led to various suggestions for regulation and moderation. However, in order to moderate, the platforms firstly have to detect it, which is not as easy as it seems, especially considering the large amount of posts that can be sent to a platform at a particular time. Consequently, some fields, such as computational linguistics and those specialising in natural language processing have invested in research, which aims to detect trolling automatically.

2.3.4. Computational Linguistic Analyses of Trolling

Computational linguists and natural language processing (NLP) researchers have focused on detecting trolling and other forms of online toxicity by developing models and training classifiers for English (e.g. Dadvar et al., 2012; Dinakar et al., 2011; Burnap and Williams, 2014; Chen et al., 2012; Xiang et al., 2012; Warner and Hirschberg, 2012; Mehdad and Tetreault, 2016; Hosseinmardi et al., 2015; Yin et al., 2009; Davidson et al., 2017; Raisi and Huang 2017; Huang and Raisi, 2018) and a variety of other languages, including Hindi (Singh et al., 2018) and Chinese (Shuang-Shuang, 2010) (see Schmidt and Wiegand, 2017*)*. The most common approach to detect them is through machine learning and this often involves having an annotated training and test dataset that is coded for a positive and a negative class (e.g. *trolling* or not *trolling, hate speech* or *not hate speech*).

The starting point for many of these studies begins with collecting corpora of normal and abusive posts, although others have used shared datasets of hate and online harassment (e.g. Golbeck et al., 2017). The abusive posts are often collected by searching for posts that contain profanity, particular hashtags, keywords or slurs with the view that these posts are likely to contain some form of abuse and hostility (Samghabadi et al., 2017; Maity et al., 2018; Abril, 2018). Because slurs, profanity and hashtags are not always abusive (Clarke and Grieve, 2017; Waseem, Thorne and Bingel, 2018; Ajayi, 2018; cf. Abril, 2018), annotators are then sourced to code the data containing these features for whether they are actually offensive or not (e.g. Golbeck et al., 2017; Davidson et al., 2017).

Having collected the data and annotated it, the subsequent step for developing a classifier is to select and extract features from each class in the training data and then use these to classify the test dataset for either trolling or not trolling based on whether these features occur in the test data and evaluate its accuracy. The selection of features in any predictive and classification task is therefore important and it is what distinguishes each study.

Words and characters are commonly used in most studies in the area of abusive language detection (e.g. Chen et al., 2012; Xu et al., 2012; Warner and Hirschberg, 2012; Burnap and Williams, 2015; Waseem and Hovy, 2016; Burnap and Williams, 2016; Hosseinmardi et al., 2015; Nobata et al., 2016), such as in the Bag-of-Words (BoW) approach (Greevy and Smeaton, 2004). BoW looks at characters and/or words (unigrams and larger sequences of n-grams) that are unique to the positive class and unique to the negative class. Models are then trained on these so that if a new message is analysed and contains a larger percentage of the words in a particular class it will be coded as that. Most studies utilising BoW have reported high predictive power with some false positives and negatives, although the performance of the model can be increased with additional features (Nobata et al., 2016). Some additional features that have been used include the frequency of punctuation, URLs, tokens with non-alpha characters in the middle, capitalised letters, and average length of words (Nobata et al., 2016), as well as tokens enriched with part-of-speech (POS) information (Xu et al., 2012), POS *n*-grams (Davidson et al., 2017), lexicons (Gitari et al. 2015), and deeper syntactic information, such as typed dependency relationships (Chen et al., 2012; Burnap and Williams, 2015; Burnap and Williams, 2016; Nobata et al., 2016).

Although models trained on BoW are relatively good at predicting and detecting abusive language, word-based models are dependent on the particular words appearing in both the training and test data. Given that in general most words occur relatively infrequently (Zipf, 1949), many models trained on specific words that occur relatively infrequently may not be as applicable or efficient to other test data sets (Eisenstein et al., 2014). Consequently, most studies using BoW have also applied some form of word generalisation, such as Brown clustering (Brown et al., 1992) (e.g. Warner and Hirschberg, 2012), Latent Dirichlet Allocation (LDA) topic modelling (Blei et al., 2003) (e.g. Zhong et al., 2016), and word, character, comment and paragraph embeddings (Mikolov et al., 2013; Le and Mikolov, 2014) (e.g. Djuric et al., 2015; Nobata et al., 2016; Pavlopoulos et al., 2017; Badjatiya et al., 2017; Mishra, Yannakoudakis and Shutova, 2018), all of which are aimed at grouping words together that occur in similar contexts in order to have a more general set of features. These techniques have been shown to increase accuracy, although there is still work to do.

Social media data does not just provide the text, but also various amounts of meta-information, and thus meta-information and extra-linguistic features have also been incorporated into models detecting abusive posts, all at varying levels of success. Some of the meta-information and extra-linguistic features that have been used include, gender and location (Waseem and Hovy, 2016), number of posts and number of replies to a post (Zhong et al., 2016), user behaviour and performance (Balci and Salah, 2015; Dadvar et al., 2013), and surrounding posts (Yin et al., 2009), including the tweet it is in reply to or the tweet it quoted (Lee, Yoon and Jung, 2018), as well as community information, such as what community the text occurred in and whether this is normal or likely to have a negative sense, called Bag-of-Communities (Chandrasekharan et al., 2017). Moreover, based on the notion that users are more likely to post abusive messages if they have written abusive messages previously, Dadvar et al. (2013) take into account the message history of a user and count the number of profane words that occur and incorporate this information into the model.

Whilst such research has focused on detecting abusive posts, other work has focused on detecting opinion conflicts (e.g. Maity et al., 2018) and problematic users. For example, Cheng et al. (2015) found differences within the posting behaviour of users who had been banned in three different discussion-based communities (Breitbart.com, CNN.com, and IGN.com). Specifically, they found that banned users who had a low post deletion rate spread their messages across a larger number of discussions, whilst banned users with high post deletion rate concentrated a lot of their messages in the discussions they contributed to.

This growing body of research gives a clear indication of each models' performance showing which combination of features is most effective for the classification task, and at what point is the model's performance impeded when more or fewer features are incorporated. Moreover, many of these studies compare the feature sets and techniques used in previous studies with the new proposed feature sets, enabling clear comparisons and baseline scores. Finally, the approach taken in many of these studies for collecting data has been fruitful in collecting large datasets of abusive language.

Despite all these positives and many more, there are a few general concerns of relevance to this dissertation. The first concern relates to the fact that abusive posts are for the most part collected if they contain profanity, hashtags or slurs. Although these posts are later annotated to ensure that they are actually abusive, these features do not necessarily have to occur or be near environments that are trolling or abusive. For example, research shows that there are more subtle forms of aggression, which are just as damaging, such as ‘othering’ – the process of negatively presenting an outgroup, often a minority group (Alorainy et al., 2018). Thus, whilst abusive posts are collected, this approach does not account for offensive or trolling posts that do not contain profanity, slurs or hashtags.

Waseem et al. (2018) introduced the next concern in their study examining the accuracy of annotations. Whilst annotators are extremely instrumental for distinguishing abusive from non-abusive posts, they do not always reach agreement because determining offence is inherently a subjective task, and this decision will be influenced by various social and cultural backgrounds (Waseem, Thorne and Bingel, 2018). They showed that many posts in an abusive language dataset (Davidson et al., 2017) were incorrectly deemed to be hate speech or offensive because they contained the n-word. They were incorrectly classified because they ended with ‘ga’ or ‘gah’ and were used amongst African Americans, and thus were not necessarily offensive or hate speech, but were being used as a part of the process of reclamation. It is therefore important for annotators to come from various social and cultural backgrounds and for disagreements between annotators to be investigated.

The final concern comes from a purely linguistic perspective. Despite most of these studies providing the precision and recall measures for various combinations of features in their classifiers on the test data sets, there is little explanation provided as to why certain features work and why others do not. Moreover, what is encompassed in the features is not necessarily referenced. For instance, assuming n-grams work in the final classifier, many studies do not detail what these n-grams are specifically. These specific realisations of features could lead to further explanations or judgments as to why the features work better than others. Moreover, the realisations of features could also be incorporated into other classifiers and evaluated on new datasets. As a result, there is limited knowledge on the major linguistic properties of hate speech, abusive language and trolling.

Moreover, apart from Singh et al. (2018) who specified different types of aggression (e.g. covert and overt), many of these studies do not appear to acknowledge or account for the variability found across the different types of abusive posts and non-abusive posts. Given that trolling is, for the most part, a linguistic act, there lacks a systematic investigation into understanding how it varies linguistically and how it varies linguistically in relation to general social media posts. In particular, still relatively little is known about the various communicative functions, communicative styles, and linguistic repertoires of trolling. Overall, the state-of-the-art models are not able to distinguish trolling from non-trolling texts with complete accuracy, which suggests that we need more understanding about the ways in which they vary. Such information about the ways in which they vary could be used and incorporated into method design for classifiers.

2.4. Summary and Research Questions

This literature review has demonstrated that trolling is a multi-faceted and multi-functional phenomenon. There is a growing body of research investigating trolling with respect to why people troll, what its effects are, how people perceive and talk about trolling, and how to detect it. This research has begun to illustrate that trolling varies considerably, detailing some of its different behaviours, types, motivations and guises and showing that it has a wide-range of communicative goals beyond provoking a response. However, there has been surprisingly little linguistic research on trolling, especially concerning the description of its linguistic repertoires and properties and how these compare to other social media posts. When so much variation exists, the question arises about whether the diversity of behaviours captured by the term ‘trolling’ is actually reflected in linguistic distinctions across numerous instances of trolling. Moreover, questions arise about how the language of trolling compares with the language of other social media posts.

Given that Twitter trolling is situated in the context of Twitter, it is possible that trolling tweets could just be drawing on the major linguistic repertoires of general tweets. It is therefore important to compare the major patterns of linguistic variation and communicative functions of trolling with those of tweets more generally. This literature review has revealed that a thorough linguistic description of the range of linguistic variation across Twitter trolling and Twitter more generally do not exist. Whilst previous literature has compared tweets in relation to the major patterns of linguistic variation of pre-internet registers, as well as other online registers, there has not yet been an investigation into how tweets vary from one tweet to the next, especially with respect to its major patterns of linguistic variation. As a result, this dissertation begins to fill these gaps by answering the following research questions:

1. What are the most dominant patterns of linguistic variation in English trolling tweets?
2. What are the most dominant patterns of linguistic variation in general English tweets?
3. How do trolling tweets compare to general English tweets?

By answering these research questions, this thesis provides the first large-scale descriptions of the range of linguistic variation and major communicative functions across general English Twitter and across Twitter trolling. Additionally, it provides the first linguistic description and comparison of trolling tweets to general tweets with respect to the major patterns of linguistic variation of general Twitter.

3. Methodology: Data Collection

This dissertation aims to provide a thorough linguistic description of Twitter trolling, not only with respect to its major communicative functions and patterns of linguistic variation, but also with respect to the major communicative functions and patterns of linguistic variation of general Twitter. It is therefore important to attempt to collect corpora that represent the linguistic distributions of these language varieties. The following section describes the method used in this dissertation for collecting the Twitter trolling corpus and the general Twitter corpus.

3.1. The Trolling Corpus

Trolling is a complex and multi-faceted phenomenon, which has been found to vary not only from platform to platform, but also within communities on specific platforms, and between individual trolls according to their varying purposes (see section 2.4). Consequently, there is not a ‘one-size fits all’ definition of trolling, as the term tends to be used as an umbrella, behavioural catch-all term (Hardaker, 2010; Phillips, 2016; Phillips and Milner, 2017). Additionally, one facet of some kinds of trolling is to deceive, convincing the community that they are genuine and that their intentions are not to purposely provoke (Donath, 1999; Hardaker, 2010). Determining someone’s intentions and whether someone is in fact trolling is impossible, unless they explicitly state so; but even then, they could be lying (Hardaker, 2015; 2018). Poe’s

Law (2005), for example, suggests that without some obvious marker (e.g. a smiley/winking) most people will not be able to tell if you are being serious or joking, even if the statement is so outrageous or extreme. Poe's Law and the deceptive nature of trolling makes identifying trolling a complex task and thus, the difficulty in analysing, understanding and classifying trolling fundamentally lies in being able to detect it in the first place.

3.1.1. Previous Approaches for Collecting Trolling

There have been four previous methods for collecting trolling, all of which are problematic. The first approach defines its characteristics and behaviours and then the researcher looks for posts that exhibit these predefined qualities (e.g. Herring et al, 2002; Abril, 2018). This method, however, is limited. First, it does not take into account the interpretation of the people within the interaction, as this may differ from the researcher's perspective (see O'Sullivan and Flanagin, 2003). Second, it does not account for the various uses of the term, especially the more modern day uses. Instead, it imposes a static definition onto a dynamic and evolving phenomenon. Third, searching for posts that display these pre-defined behaviours in an unbiased way on a platform like Twitter is difficult. Finally, given Poe's Law (2005) and the deceptive nature of some cases of trolling, it is impossible to know if the post identified as trolling is even such.

As an extension of the first approach, the second method collects trolling and abusive posts by using particular linguistic markers, such as swearing, name-calling, slurs and specific hashtags (e.g. Cho and Kwon, 2015; Synnott et al., 2017; Golbeck et al., 2017; Abril, 2018). Trolling, however, does not necessarily have to include name-calling, slurs, profanity or hashtags. For example, many of the trolling tweets in Clarke (2018) do not contain any of these features. More important, posts that have these features are not always abusive. Profanity, for example, can be used for amplification, and name-calling and slurs can be discussed amongst friends in a non-targeted way (Clarke and Grieve, 2017). Ajayi (2018), for example, showed how superficially abusive comments were actually used amongst Yoruba youths in Nigeria to commend, salute and praise each other.

The third approach gathers posts from self-identifying trolls (those that call themselves such) (e.g. Phillips, 2016). While this does find instances of trolling, self-identifying trolls are just one sub-culture of trolls (Phillips, 2016), as not all trolls are this forthright in marking their identity, meaning that the study would be limited to this particular group of trolls. From previous research (e.g. Donath, 1999; Hardaker, 2010), it is possible to assume that self-identifying trolls are far less likely to display behaviours intending to deceive their target of their trolling intentions. Therefore, these behaviours and functions of trolling may not be represented in a study, which only collects trolling posts from those who self-identify.

The fourth method is the least problematic, which involves using the perceptions of others and subsequently involves extracting posts that have been accused by other people to be trolling (Mihaylov and Nakov, 2016; Synnott et al., 2017; Clarke, 2018; Clarke, 2019). The most inclusive approach to this final method involves searching for the term ‘troll’ or some variation of it (e.g. trolling), and then these posts require manual examination to identify if they are an accusation (Mihaylov and Nakov, 2016; Clarke, 2018). For example, Mihaylov and Nakov (2016) crawled a Bulgarian community forum and extracted posts that were accusing others of trolling. Specifically, they considered a comment as a potential accusation if it was replying to a comment and if it contained the word *troll* or its Bulgarian equivalent. These posts were subsequently checked for whether they were accusations or not. The posts that the accusations were replying to were subsequently extracted and analysed.

This method, however, is still problematic in the sense that the accused post may not actually be trolling. Additionally, some kinds of trolling posts may be more likely to receive accusations, whereas highly deceptive instances of trolling may go undetected, meaning that the corpus may not be representative of all cases of trolling. Moreover, this approach is labour-intensive because *troll* appears frequently on social media and it is not always used to accuse another person. Clarke (2019) sought to speed up this process by selecting and searching for the imperative ‘stop trolling’ because, as a directive to stop the current behaviour, it is responsive, suggesting that the post before this instruction was the post perceived to be trolling. Using the search string ‘stop trolling’ in the Twitter search tool, Clarke (2019) examined the tweets containing the phrase for accusations (e.g. *I wish everyone would stop trolling* = non-accusatory vs. *@username Stop trolling you buffoon!* = accusatory). Those that were accusations and were in reply to another tweet were clicked on to reveal this replied to trolling tweet. Additionally, Clarke (2019) also selected tweets containing the accusation ‘stop trolling’ that quoted another user’s tweet. The quoted tweet alongside the accusation ‘stop trolling' suggested that this quoted post was also perceived to be trolling. The following examples are taken from Clarke (2019) to demonstrate this. For Example A, the first tweet from @username1 would be extracted, whereas for Example B, the quoted tweet from @username5, marked by the square brackets, would be collected.

<u>Examples from Clarke (2019):</u>

A. @username1: @username2 Isn't it time you step aside and let someone who knows what they are doing, run the country?
   @username3: @username1 Stop trolling [name]!
B. @username4: Stop trolling [@username5: I think White women understand black men the best]

Although this is a principled way of speeding up the task of manually examining each post containing *troll* for an accusation, it limits the kinds of trolling posts that can be collected - because not all trolling posts are responded to in this way (e.g. *you are a troll*, *troll level 100*). 'Stop trolling' as an accusation suggests that the individual is a repeat trolling offender, and thus this approach might not have been inclusive of the trolling posts of first-time and/or deceptive trolls. Additionally, it might have only obtained posts midway through an interaction, as opposed to all of the turns of the troll.

The present dissertation followed the method of Mihaylov and Nakov (2016). Specifically, tweets labelled as trolling (or some variation of *troll*) by other Twitter users were collected for analysis. This approach is the most inclusive in that it does not limit the range of behaviours that trolling can display via some pre-defined list or set of linguistic markers. This approach also takes into account the perception of the people within the interaction, although given Poe's Law this is still questionable, as accusations could be sarcastic. Whilst this approach is the least problematic, it is still limited. In particular, it does not account for undetected trolling posts or posts that were not accused. Additionally, it does not account for trolling posts that were accused without using 'troll' (e.g. *I am not going to feed you*). Moreover, there is the possibility that the collected posts are not all the turns of the troll but actually come midway through the interaction. In short, the trolling corpus collected in this dissertation represents this sampling criteria - trolling posts that were labelled as trolling - at the time of data collection. Other instances of trolling were not considered.

3.1.2. Collecting Twitter Trolling

Based on the corpus-based linguistic understanding that words gain meaning through their use, I adopt Mihaylov and Nakov's (2016: 403) operationalisation of *troll*, which is "somebody who was called such by other people" and I use the perceptions of others to identify examples. In other words, if something is labelled as trolling, then I take it to be such because each use of the term contributes to its meaning. Thus, trolling posts were collected by firstly collecting posts that contained the word 'troll'. Then, if these posts that contained 'troll' were in reply to another tweet, this replied to tweet was also collected. The post containing the word 'troll' that was collected first was subsequently examined for whether it was an accusation to the person they were replying to or not. Only those posts that were clear accusations to the person that they were replying to, and which had the replied to tweet available were retained (see Table 1). This same approach was used in Clarke (2018).

This method for data collection was implemented in R using the 'twitteR' package (Gentry 2016). This package requires the researcher to have a Twitter account and API key, which can be obtained by following the instructions on Twitter's developer website (Twitter 2017). This key permits access to the public API stream of tweets. For research purposes, I created a Twitter account, which does not follow anyone, and I also created an API key connected to this account, which was used to collect tweets from the public stream. Subsequently, the "searchstring()" function in the twitteR package was used to collect tweets. This part of the programme extracts tweets from the public API that include a particular search string (searchString = "insert text you want to search for"), that were posted from the moment of searching up to 6–9 days before. There are specific arguments within this function to collect tweets to suit one's analysis, including the maximum amount of Tweets to return ("*n*= 25" *by default*), and the language of the tweets ("lang = en", *for English only*). There are particular limits on how many tweets a user can collect, meaning a large number will not always be permitted.

For the present analysis, the search string selected was 'troll', the language was restricted to only English tweets, and the maximum amount of tweets selected was 10,000. Each tweet containing this search string was returned with various metadata, including whether it was in reply to another user and the screen name of that user, and the ID of the post (termed 'status ID') that it was in reply to (if applicable). All the status IDs that the tweets containing ‘troll’ were in reply to were extracted from this file and then using the IDs, these posts were collected using the function ‘lookup_statuses()’.

Table 1: Manual examination of Tweets containing "troll" for an accusation (adapted from Clarke, 2018)

	Potential Accusation containing “troll”	Accusation (Yes or No)	Extracted post via statusID that the “troll” Tweet was in reply to
1.	@usrlogin @GoAngelo @seanhannity You better go back to nazi troll school. Your memes are old.	Yes	@sparkman92 @GoAngelo @seanhannity https://t.co/K55hzjjspU
2.	@0luwapemi She’s a stupid troll, don’t give her the attention she is craving	No	A grown woman tweeted this This was after complementing the abuser's penis https://t.co/bnWbccDt6o
3.	@eugenegu @donmoyn @DonaldJTrumpJr @wikileaks You're apparently not a scientist, just a Trump hating troll.	Yes	@donmoyn @DonaldJTrumpJr @wikileaks This Wikileaks scandal will go down in history as even bigger than Watergate.
4.	@larryasselin @RealMattCouch @seanhannity Aww larry is a troll	Yes	@RealMattCouch@seanhannity Who gives a shit. He lies every time he opens his mouth

Sometimes the posts were not available, potentially because it was deleted or that particular user has a private or suspended account. In this circumstance, ‘NA’ was returned. These posts were then aligned with their counterpart tweet containing the word “troll”. The tweet containing the word “troll” was then manually examined for an accusation. Table 1 illustrates this process. In some cases, accusations were present, but they were about a person external to the conversation, such as example 2 in Table 1, meaning that the post to which the person was replying was not the trolling post. These posts were ignored and not collected. In other cases, direct accusations were determined by using the name of the poster and the username that the person was replying to/accusing, like example 4 in Table 1, where an individual called ‘larry’ is described as being a troll, and this is in reply to a username called @larryasselin.

Table 2: The Twitter Trolling corpus: Dates of collection and the frequency of trolling posts collected

Data Collection Date	Trolling posts	Cumulative Frequency
13-11-2017	853	853
27-11-2017	973	1826
28-12-2017	867	2693
28-01-2018	644	3337
28-02-2018	917	4254
Repetitions	72	**4182**

Those tweets that were responded to with an accusation of being a troll were retained. The first iteration of this process of collecting posts containing the word ‘troll’ and the replied-to tweets returned 853 trolling posts. Although 853 texts is still a substantial amount of texts, especially for a Multi-Dimensional Analysis, this process was repeated four more times. Table 2 details the date of collection and the amount of posts returned from each round of 10,000 potential troll tweets. Finally, those tweets that were responded to with an accusation of being a troll were compiled into one text file and any repetitions were removed. The total corpus contains 4,182 tweets, posted from 13th November 2017 to 28th February 2018, totalling 102,191 word tokens with a mean tweet length of 24 words long. The median tweet length is 17 words long with an interquartile range of 8 to 30 words. Information on the number of users is not available as the text of the troll tweet was only extracted, although it is assumed that some trolls will be repeat offenders.

3.2. The General Twitter Corpus

To compare trolling tweets to general tweets, a general Twitter corpus was collected. This corpus serves as a baseline for trolling tweets to be compared. This corpus was collected after the trolling corpus to ensure that there was no overlap of texts between the two corpora. This corpus was collected using a different R package called ‘rtweet’ (Kearney, 2018). The reason for this difference was because at the time of collection, the ‘twitteR’ package was still based on the 140-character limit. On November 7th 2017 this character restriction on tweets was increased to 280 characters. This meant that the tweets that exceeded 140 characters were returned up to 140 characters with a URL at the end, which could be entered into a web browser to access the full tweet. Whilst this was not a problem for the trolling corpus because it required manual examination, this was not ideal for the general Twitter corpus considering the amount of tweets that were collected. As a result, I selected the ‘rtweet’ package because Kearney (2018) the developer of ‘rtweet’ had updated the package software so that the full tweet was returned. This updated version can be accessed via the development version of the package (https://github.com/mkearney/rtweet).

To collect the general Twitter corpus, I used the 'stream_tweets()' function, which returns a small random sample of public tweets (around 1% at any particular time). The aim was to collect tweets for a week using this programme, beginning on 22nd June 2018. Despite organising for the programme to run for a week, it stopped due to an Internet connection drop, working only for approximately four hours. In these four hours, 123,330 tweets had been collected. After removing non-English tweets and retweets, the final corpus contained 13,879 original (i.e. not reposted) English tweets, totalling 230,748 word tokens. The mean tweet length of this corpus of tweets is 17 words. The median tweet length is 11 words long with an interquartile range of 6 to 19 word tokens.

Importantly, there is no doubt that this corpus contains trolling tweets. The purpose is not to have a trolling corpus and a non-trolling corpus. Rather, the aim is to have a corpus that is representative of the range of linguistic distributions of English tweets generally. Given that trolling has become a major part of Twitter, it would be wrong to exclude it from a corpus that is aimed at representing general Twitter. Nevertheless, the frequency of trolling posts in this corpus is not known.

This corpus is limited in that it does not span a large time frame, meaning that there is the possibility that the results of the analysis may not be the same with a corpus of tweets collected over a longer time period. Additionally, the corpus is a comparatively small corpus of tweets. The size of this corpus may influence the range of linguistic distributions observed, where a larger corpus may contain more variation. Whilst research has suggested that quantity can often trump quality with respect to corpus design (e.g. Morstatter et al., 2013), there have also been suggestions that smaller corpora are adequate for investigations of high frequency features (Biber, 1993; Egbert, 2019). Nevertheless, it is important that the corpus is evaluated for its representativeness. Consequently, this corpus and the trolling corpus were evaluated for their representativeness in terms of representing the range of linguistic variation across the varieties (see section 4.8).

Specifically, smaller random samples of the corpora were assessed for their linguistic distributions. These linguistic distributions across the smaller random samples were correlated to the linguistic distributions of other smaller random samples of the particular corpus. From this analysis, it was found that the linguistic distributions in these smaller samples were strongly correlated to each other, indicating that the data had reached a point of saturation at these smaller sample sizes, where additional data did not reveal any more or any new linguistic distributions. The method for assessing the representativeness of the corpora and results are described in more detail in section 4.8. In short, the results indicate that the corpus is representative of the range of linguistic distributions on Twitter from this particular time period. However, it is possible that over a longer time frame these could change.

4. Methodology: Data Analysis

The approach selected to identify and compare the range of linguistic variation across the two corpora is Multi-Dimensional Analysis (MDA) because it not only enables the identification of the major patterns of linguistic variation across a corpus of texts, but it is also an approach for systematically comparing new texts to existing patterns of variation. The following chapter describes the approach, its theoretical underpinnings, and the results from Biber (1988). Subsequently, previous research employing MDA is reviewed and the limitations of the approach are described, paying specific attention to two problems with applying MDA to tweets. The solutions to these problems that this dissertation offers and applies to the corpora are then explained. Finally, the method used to assess the representativeness of the corpora is also described.

4.1. Biber’s Multi-Dimensional Analysis

Multi-Dimensional Analysis (MDA) is a corpus-based, data driven approach, pioneered by Biber (1984, 1985, 1986, 1988), which combines computational approaches, linguistic and statistical analysis to systematically investigate and describe the overall parameters of linguistic variation within a given domain. In these original studies, Biber set out to examine the variation across spoken and written English. He was influenced by theoretical discussions (Ervin-Tripp, 1972; Hymes, 1974; Brown and Fraser, 1979) that emphasised the importance of linguistic co-occurrence when analysing registers and the functional differences between them (Biber, 2019). These discussions highlighted that linguistic features vary because they serve functions, and thus there is a relationship between communicative function and patterns of linguistic variation. For example, Brown and Fraser (1979) argued that exploring the co-occurrence of sets of linguistic features enables systematic functional variations of texts to be taken into account.

In addition to these theoretical discussions, Biber was also influenced conceptually and methodologically by Carroll (1960), who subjected the frequency counts of the occurrence of 39 linguistic features in 150 texts to factor analysis with “the aid of high-speed electronic computing machines” (1960: 280), revealing six vectors of prose style (Biber, 2014: XXXI). Crucially, this study revealed that linguistic variation is amenable to statistical analysis, specifically factor analysis, as it can be used to reveal sets of linguistic features whose frequencies are correlated across several texts in a corpus of texts. Thus, in light of the theoretical discussions of linguistic co-occurrence and Carroll’s (1960) research, MDA is aimed at identifying the most common patterns of co-occurring linguistic features across a corpus of texts by subjecting the normalised frequencies of numerous lexical and grammatical features in the texts of a corpus to factor analysis.

It achieves this by firstly tagging each text for a variety of different lexico-grammatical features, ranging from grammatical parts-of-speech (e.g. prepositions, pronouns), and semantic categories for verbs (e.g. private and public verbs) to longer syntactic constructions like relative clauses. After tagging, the next step is to measure the relative frequency of each linguistic feature in each text across the corpus of texts and record these frequencies in a continuous data matrix, where each row represents a text and each column is a linguistic feature. Each cell in a row shows the frequency of a particular linguistic feature in a text relative to the length of the text in words (e.g. 15 times per 1,000 words). Having recorded this information, the next step involves subjecting this data matrix to factor analysis.

Factor analysis is a multivariate statistical method used to find underlying or latent variables by finding variation in observed and correlated variables. Factor analysis produces a series of factors or dimensions that represent the shared patterns of variation of the variables across the individuals, with each subsequent dimension representing another distinct pattern of covariation. When factor analysis is applied in MDA to the data matrix of the relative frequencies of numerous linguistic features across the texts of a corpus, it reveals a smaller number of factors or dimensions, representing the most frequent patterns of co-occurring linguistic features.

Specifically, each factor has a weighted combination of all the linguistic features, where each linguistic feature has some weight for each factor. Each linguistic feature's weight is its loading and loadings range from -1 to +1 for each factor, indicating the amount of shared variance with the total pool of variance. The strength of this loading represents how associated it is to the factor. Loadings that are closer to 0 tend to be ignored, as they are not really influencing the factor. Loadings that are closer to -1 or +1 are given prominence, as the variables assigned high weights are relevant for the factor. In other words, these loadings show which linguistic features tend to statistically co-occur with other linguistic features most frequently. Each subsequent dimension reveals the next most common pattern of linguistic features that tend to co-occur in the texts by assigning each linguistic feature a different weight. Most features load strongly on the early factors, and also they can load strongly on more than one factor. To make the results more interpretable, a process of factor rotation is completed, usually Promax, which makes it easier to identify a variable, as being associated with a single factor.

Following this quantitative statistical component, the next step in MDA is interpretative and functional, where the dimensions of aggregated linguistic co-occurrence patterns, and their realisation in the texts most associated to them are interpreted as latent causes of linguistic variation in the corpus. Specifically, the analyst makes a decision about what linguistic features to take into account for each dimension based on the strength of the loadings. Most MDA studies consider factor loadings above 0.3 as strong, although this varies. Additionally, the analyst computes factor scores for each text, which indicate how associated each text is to the particular patterns of linguistic co-occurrence captured by the dimensions. Based on the notion of linguistic co-occurrence, the linguistic features with strong loadings for each dimension returned by the factor analysis are subsequently interpreted along with the texts displaying these patterns (i.e. those with high factor scores) for the underlying communicative function. The interpretations of the dimensions need to capture the function that is shared by the co-occurring features.

Importantly, each dimension consists of a positive pole and a negative pole. Each pole is associated with a set of co-occurring linguistic features that are in complementary distribution with the other set on the opposite pole. Each dimension is conceptualised as a continuum of variation, where the set of co-occurring linguistic features on one side is more associated with the function, whereas the other set of co-occurring linguistic features is less associated with the function. Thus, when it comes to interpreting the dimension, the label assigned to the dimension must capture the function that explains the difference between the two sets of co-occurring features. For example, the degree of interactivity - the features on one side are more interactive and the features on the other side are less interactive.

Using this method, in one of the first large-scale studies examining the variation across spoken and written language, Biber (1988) revealed 6 dimensions of linguistic variation, although only 5 tend to be reported. The interpretative labels assigned to these dimensions and brief descriptions of the patterns of variation are presented below.

Dimension 1. Informational versus Involved Production: the linguistic features with negative loadings mark high informational density and specific informational content (nouns, prepositions, attributive adjectives, type/token ratio, word length, etc.), whereas the linguistic features with positive loadings are less specific and mark generalised content. Additionally, the features with positive loadings are more affective and interactive (e.g. first and second person pronouns, private verbs, demonstrative pronouns, contractions, WH questions, etc.). This dimension not only reflects the primary purpose of the author/speaker, but also the production circumstances - informational texts tend to have complex structures and dense noun phrase modification and these are enabled when there is time to edit and select precise lexis, as opposed to interactive texts, which are influenced by real-time constraints, leading to less precise lexis, more pronouns and contracted forms.

Dimension 2. Narrative versus Non-narrative Concerns: the linguistic features with positive loadings mark narrative concerns in that they function to mark past time, third person animate referents, reported speech, and depictive discourse (e.g. third person pronouns, past tense verbs, perfect aspect, public verbs, etc.), whereas the linguistic features with negative loadings are non-narrative as they function to mark immediate time and a more frequent elaboration of nominal referents (e.g. present tense verbs, attributive adjectives, past participial WHIZ deletions).

Dimension 3. Explicit versus Situation Dependent Reference: the linguistic features with positive loadings mark exophoric references that are highly explicit and context-independent (e.g. WH relative clauses on object and subject positions, pied-piping constructions, nominalisation and phrasal coordination), whereas the features with negative loadings mark endophoric, non-specific and situation-dependent references (e.g. time and place adverbials, adverbs).

Dimension 4. Overt Expression of Persuasion: there are only linguistic features with positive loadings and these features function to persuade the addressee, either by indicating the speaker’s own point of view, or by assessing the advantages or the likelihood of an event (e.g. infinitives, prediction modals, suasive verbs, conditional subordination, necessity modals, split auxiliaries, possibility modals).

Dimension 5. Abstract versus Non-abstract Information: the linguistic features with positive loadings are used to indicate informational discourse that is abstract, formal and technical (e.g. conjuncts, passives, adverbial subordinators, past participial clauses and WHIZ deletions, predicative adjectives), whereas the negative loadings are used to mark other kinds of discourse (e.g. type/token ration).

Dimension 6. On-line Informational Elaboration: the linguistic features with positive loadings function to mark fragmented informational elaboration that is relatively spontaneously produced, especially under strict real-time constraints (e.g. THAT clauses as verb and adjective complements, THAT relative clauses and WH relative clauses on object position, final preposition, existential THERE, demonstrative pronouns), whereas the linguistic features with negative loadings are associated with informational integration (e.g. phrasal coordination).

4.2. Previous MDA studies

Whilst MDA began as an approach to investigate the variation across spoken and written language in English, it has since been used to investigate language in use in numerous other languages and specialised discourse domains. For example, it has been used to examine variation across the registers in languages, such as Nukulaelae Tuvaluan (Besnier, 1988), Somali (Biber and Hared, 1992; 1994), Korean (Kim and Biber, 1994), Taiwanese (Jang, 1998), Dutch (Grieve et al., 2017b), Brazilian Portuguese (Sardinha, Kauffman and Acunzo, 2014), Gaelic (Lamb, 2008), Spanish (Biber et al., 2006; Biber and Tracy-Ventura, 2007; Parodi, 2007; Asención-Delaney, 2014), Russian (Katinskaya and Sharoff, 2015) and World(-wide) English(es) (Xiao, 2009; Bohmann, 2017).

Additionally, MDA has been applied to specific domains, such as the spoken and written registers in elementary school (Reppen, 1994; 2001), and university (Biber, 2006), as well as the variation in academic research articles (Gray, 2011; Gray, 2013; Thompson et al., 2017), moves in science research articles (Kanoksilapatham, 2007; Biber and Jones, 2005), and in spoken and written exam responses by English as a Second Language learners (Biber and Gray, 2013). Additionally MDA has been applied to spoken registers like job interviews (White, 1994), conversation (Biber, 2004), call centre discourse (Friginal, 2009), and cases of cross-talk (Connor-Linton, 1999), as well as written registers like legal texts (Goźdź-Roszkowski, 2011), and newspaper editorials (Westin and Geisler, 2002). MDA has been applied to registers across specific time periods, such as the written and speech-based registers in the 18th century (Biber, 2001), fictional novels in the 19th century (Egbert, 2012), philosophical transactions of the Royal Society of London from 1675-1975 (Atkinson, 1999), North American movies between 1930-2010 (Pinto, 2014), North American and British pop songs from 1940-2009 (Bértoli-Dutra, 2014), TIME magazine cover stories from 1923-2011 (Condi de Souza, 2014) and Donald Trump’s tweets from 2009 to 2018 (Clarke and Grieve, 2019).

MDA has also been used to investigate particular phenomena, such as the variation in the use of metadiscourse markers in some spoken (Zhang et al. 2017) and written registers (Zhang, 2016), essay writing quality (Crossley, Allen and McNamara, 2014), colloquialisation in original films and remakes

(Zago, 2017), the similarities and differences between natural and television versions of a particular register, for instance comparing television dialogue on sitcom Friends with natural conversation (Quaglio, 2009), and comparing high profile criminal trials with television series courtroom trials (Chen, 2018), and it has even been used to investigate regional variation in American English (Grieve, 2014a, 2014b, 2016).

Related to the present thesis, MDA has also been applied to online communication, such as bulletin boards (Collot and Belmore, 1996), websites (Biber and Kurjian, 2007), web-blogs (e.g. Grieve et al., 2011; Hardy and Friginal, 2012), some kinds of web registers (Titak and Roberson, 2013; Passonneau et al., 2014; Sardinha, 2018), Nordic Twitter Englishes (Coats, 2016), Russian web registers (Katinskaya and Sharoff, 2015), and the searchable web (Biber and Egbert, 2016; Biber, Egbert and Zhang, 2018).

These show that MDA is a powerful method for investigating large corpora of language in use. Additionally, it permits rich descriptions of language in use, documenting how language users consistently make certain language selections in particular contexts and situations for particular purposes. Given that variation is inherent in language, these studies all reveal unique dimensions of linguistic variation of the particular languages, discourse domains, registers and texts under investigation. However, one of the most important findings of the majority of these studies is that Biber's (1988) first (Informational versus Involved Production) dimension, and often second (Narrative versus Non-narrative) dimension, have been found consistently across various languages and discourse domains, suggesting that these are potentially universal dimensions (Biber, 2014; Biber, 2019). Surprisingly, the first dimension, which typically opposes a formal and literate kind of style with a more oral and informal style has even been discovered in studies focusing exclusively on spoken registers (e.g. Biber, 2004) and also on written registers (e.g. Gray, 2013).

In addition to running full MDA, as described here, it is also possible to compare new varieties of language to existing dimensions of linguistic variation by projecting them onto the existing dimensions. This version of MDA does not aim to identify new dimensions of variation, but rather compares and maps on different language varieties along existing dimensions of variation (Sardinha, 2014). It does this by similarly tagging each new text for the linguistic features used in the original MDA. The relative frequencies of these features are computed and then based on the mean and standard deviation scores of the original analysis, factor scores of the new texts and registers are calculated (Sardinha, 2014).

The majority of studies in this area compare English registers to the dimensions of spoken and written English in Biber (1988). Examples of these other English registers and varieties include malicious forensic texts (Nini, 2017; 2019), written English for Academic Purposes essays (Crosthwaite, 2016), learner English (van Rooy and Terblanche, 2009), first and second language writing development of elementary students (Reppen, 2007), Internet registers (Sardinha, 2014), synchronous and super-synchronous CMC (Jonsson, 2015), as well as East African English (van Rooy, Terblanche, Haase and Schmied, 2010), and translated and non-native indigenised varieties of English (Kruger and van Rooy, 2016). All of these studies enable rich comparative descriptions of registers that were not included in Biber's

(1988) study with the ones that were (see Sardinha et al., 2019). Overall, MDA offers a quantitative approach for not only identifying the major patterns of linguistic variation across a corpus of texts, but also an approach for systematically comparing new texts to existing patterns of variation.

4.3. Limitations

Despite the method's power in enabling rich descriptions of language variation, it is necessary to acknowledge that Biber's approach has often been met with criticisms. The first limitation of studies employing MDA concerns the feature set. The ability to identify the most important underlying patterns of variation depends largely on the decision about what variables (i.e. linguistic features) are included in the first place (Grieve, 2016; McEnery et al. 2006). Whilst MDA aims to be bottom-up and all-inclusive of a range of lexico-grammatical features, the feature set is nonetheless selected and defined by the author. Biber's (1988) feature set relied "on form-to-function correlations established in micro-analyses" (Biber, 1988: 63) of spoken and written English. Many studies employing MDA have adopted this feature set used in Biber (1988), but have failed to incorporate more specific features relevant to the discourse domain under investigation. Because of this the features selected in studies employing MDA are often critiqued, with recommendations to be as inclusive as possible (McEnery et al., 2006).

Whilst it is true that some studies using MDA have not been as inclusive as possible in their selection of features, the inclusion of features is very often dependent on the ability to design algorithms to detect their occurrence computationally, especially in large corpora of texts. For instance, it is considerably complex (if not impossible) to design an algorithm to distinguish all gerund forms without checking them by hand, meaning that only some types of gerunds are identified automatically or the researcher has to spend considerable time checking. Thus, rather than being a general limitation of Biber's MDA, it is a general problem of tagging and the application of Biber's MDA, where researchers have not always attempted to include specific features.

The next limitation concerns how deeply involved the researcher is during the interpretation of the dimensions (Sardinha and Pinto, 2014), meaning that criticisms have also concerned the subjectivity and clarity of the functional interpretations (Santini, 2005). It is important to remember that the labels assigned to the dimensions of linguistic variation are interpretations, and therefore they do not influence the result of the statistical analysis. It is possible for others to disagree with the naming. Nevertheless, the linguistic co-occurrence patterns identified through the factor analysis are statistically sound relative to the corpus. Recent research, however, has focused on developing methods to validate the dimensions of linguistic variation and the interpretations (e.g. Pavalanathan et al., 2017). Although this has been a criticism attributed to MDA, it is nevertheless a critique of almost all quantitative linguistic studies, as it is the linguist's task to interpret and explain the patterns. For instance, in a collocation analysis, the linguist must group and interpret patterns of collocates. The only linguistic studies which are not met with this criticism are some computational approaches, especially in predictive tasks, as they do not seek to explain the patterns, although this is arguably a limitation of computational approaches. Thus, whilst the criticism that the labels assigned to the dimensions in MDA are subjective is true, this is not an exclusive criticism of MDA, but rather it is a criticism of quantitative linguistic studies more generally. Moreover, it is a necessary part of most linguistic studies, where the linguist attempts to make sense of the data and patterns observed.

Another limitation is that the method is aimed at identifying the most important/frequent patterns of variation, which arguably ignores or overlooks minor parameters of variation (Biber, 1988). It is important to note that MDA cannot do everything and it does not set out to do everything. It is focused on the more general and most frequent patterns of variation, rather than the more nuanced kinds.

Whilst these are general limitations of the approach, there are also two problems that are directly related to applying MDA to short texts like tweets, which are the main aims of this dissertation. As a result, these problems and the solutions that this dissertation offers are presented separately below.

4.4. Problem 1: Tagging and the Feature Set

The first problem with applying the standard form of MDA to texts, such as tweets is that tweets can include lots of non-standard spelling and grammar. As a result, most taggers are not well suited to accurately tagging tweets. Because the analysis is based on the frequencies of grammatical features, inaccurate tagging can therefore lead to inaccurate co-occurrence patterns being revealed. Thus, it is necessary to tag the texts in the most accurate way possible. Different solutions to this have been offered in previous MDA studies that have investigated tweets. For example, Passonneau et al. (2014) in their MDA study of tweets and other genres used the manually annotated sub-corpus (MASC), and thus did not necessarily have to worry about inaccurate tagging, as the tagging of the texts is manual and is checked by various other annotators. Nevertheless, the feature set (all the annotations in the MASC) is considerably smaller and less detailed than that used in traditional MDA research.

Titak and Roberson (2013) and Friginal, Waugh and Titak (2018) used a more traditional feature set by tagging the tweets using the Biber Tagger, although they do not specify how accurate it was or whether the tags were checked and amended appropriately. Despite having a much larger and MDA-specific feature set than Passonneau et al. (2014), these studies that used the traditional MDA feature set on tweets have not incorporated additional features relevant to the discourse domain under examination. For example, it is standard practice in full MDA studies to incorporate features into the feature set to represent the data under investigation (e.g. Grieve et al., 2010), which Passonneau et al. (2014) and Coats (2016) did for their analysis of tweets. For instance, Coats (2016) conducted an MDA of Finnish English Tweets and global English Tweets, and included hashtags, @-mentions, emojis/emoticons and URLs, among other features in his comparison of tweets, and Passonneau et al. (2014) included different types of named-entities. Therefore, it can be argued that Titak and Roberson's (2013) and Friginal, Waugh and Titak's (2018) functional linguistic descriptions of groups of tweets are not as informative as they could be with a more detailed feature set that is specific to the language under investigation.

There has been work in the NLP field to deal with non-standard text, especially that found on social media. Techniques such as normalisation and domain adaptation have been proposed. Normalisation involves adapting the text to fit the tools (e.g. the conversion of nonstandard language to standard) (e.g. Liu, Weng and Jiang, 2012), whereas domain adaptation involves adapting the tools to fit the text (e.g. creating new annotation schemes for social media data) (e.g. Gimpel et al., 2011). It is important to note that the theories behind non-standard language use that these approaches are based on have been criticised, as they fail to acknowledge that non-standard language use can be purposeful, such as to mark one's identity (see Eisenstein, 2013).

One example of domain adaptation has been the development of the Twitter Tagger (Gimpel et al., 2011; Owoputi et al., 2012; Owoputi et al., 2013), which tags tweets for 25 features, including traditional part-of-speech tags (e.g. nouns, verbs, determiners and adjectives), Twitter and online-specific features, (e.g. hashtags, URLs and @-mentions), as well as combined part-of-speech tags (e.g. nominal + verb = *I'm*, proper noun + possessive = *Donald Trump's* hair, nominal + possessive = The *cat's* dinner*,* proper noun + verb = *Donald Trump's* got no hair). The tagger can be trained on various models to lead to different tagsets, although these need to be specified, otherwise the default is used. For example, there is a model that incorporates the Penn Treebank part-of-speech tag set, and this model was used by Coats (2016) in his MDA study of tweets. Like Passonneau et al.

(2014), the feature set used in Coats (2016) is not as detailed as traditional MDA studies, although it does incorporate features related to Twitter. Overall, previous MDA studies of Twitter have not incorporated both the MDA feature set *and* additional features related to the discourse under examination.

4.5. The Solution: Multi-Dimensional Analysis Twitter Tagger

In this dissertation each tweet in the two corpora needed to be tagged for various lexical and grammatical features, including those in the standard MDA feature set, as well as those related to CMC and Twitter specifically. To do this, the present dissertation used the 0.3.2 version of the Twitter tagger developed by Gimpel et al. (2011) and Owoputi et al., (2012) with the original/default tags (see Gimpel et al., 2011) in the Computational Natural Language Learning (CoNLL) output format. Specifically, this version of the tagger requires that each tweet be on a new line in a single text file. For example, consider Examples 1 to 4 below, which are four trolling tweets in their bare form separated by a newline.

Demonstration of tagging

1. @username Sit down honey.
2. @username Grow up lady.
3. @username This is gold
4. @username that’s not accurate

The CoNLL output format means that the tagger tokenized each tweet in the files of trolling and general English tweets, and each token was placed on a new line and subsequently assigned one of 25 tags separated by a tab and a confidence measure between 0 and 1 also separated by a tab, where 1 indicates 100 percent confidence in the tag assigned to the token. Examples 1i to 4i represent this output on examples 1 to 4.

1i.	@username	@	0.9989
	Sit	V	0.9764
	down	T	0.8888
	honey	N	0.9676
	.	,	0.9985
2i.	@username	@	0.9989
	Grow	V	0.9938
	up	T	0.9712
	lady	N	0.9836
	.	,	0.9985
3i.	@username	@	0.9990
	This	O	0.9491
	is	V	0.9972
	gold	A	0.9007
4i.	@username	@	0.9989
	that’s	L	0.9974
	not	R	0.9983
	accurate	A	0.9524

Using regular expressions, this file was cleaned to remove the confidence measures and attach the tag to the token with an underscore (e.g. went_V), as can be observed in examples (1ii)-(4ii).

1ii. @username_@

Sit_V

down_T

honey_N

._,

2ii. @username_@

Grow_V

up_T

lady_N

._,

3ii. @username_@

This_O

is_V

gold_A

4ii. @username_@

that's_L

not_R

accurate_A

Finally, each tagged token in a tweet was combined on a single line with each tweet on a separate line, depicted in examples (1iii)-(4iii).

1iii. @username_@ Sit_V down_T honey_N ._,

2iii. @username_@ Grow_V up_T lady_N ._, 3iii. @username_@ This_O is_V gold_A

4iii. @username_@ that’s_L not_R accurate_A

Whilst this tagger is able to account for the noisiness of Twitter spelling, as well as tag for features specific to Twitter, for example URLs, hashtags and mentioning with a high accuracy (almost 90%), the overall feature set is nowhere near as detailed as the one used in standard MDA (see Biber, 1988). Additionally, some features are conflated into one tag, such as prepositions and subordinators. Consequently, I developed a Twitter-specific MDA tagger, called MDATT (Multi-Dimensional Analysis Twitter Tagger), which is based on the output of the Twitter tagger and tags each Tweet for a 124 linguistic features, according to the MDA feature set, as well as other features related to CMC. Table 3 presents each feature and gives a brief description of it and an example, and its functional association. Non-standard spellings were also included in the tagger because these are common on Twitter.

The MDATT is written in Perl. Firstly, it begins by taking the original tagging output of Gimpel et al.’s (2011) Twitter tagger and reassigns tags to add more detail. For example, it separates subordinators and prepositions, which are assigned the same grammatical tag _P by the Gimpel tagger. Additionally, it assigns an additional tag with a tilde ‘~’ and underscore ‘_’ in cases where adjectives and nouns have been tagged as a verb (e.g. a_D left_A *leaning_V* site_N to a_D left_A *leaning_V_ADJ* site_N). Secondly, it assigns a basic part of speech tag to all features, which is often the same that has been assigned by the Gimpel et al. (2011) tagger, albeit the more detailed labels and fixed errors from stage 1 (e.g. a_D_DET left_A_ADJ *leaning_V_ADJ* site_N_NOUN).

Following this, the tagger incorporates the rule-based algorithms described in the appendix of Biber (1988) and the particular lexicons specified in Quirk et al. (1985) and Biber et al. (1999), adapted to accommodate this tagger and common spelling variations found across CMC (e.g. *gonna* for *going to, dis* for *this*). Specifically, rule-based algorithms are used to tag features by using specific words (sometimes from particular lexicons), tags and sequences of words and tags, taking into account the various spelling variations inherent in CMC. For example, infinitives are tagged if the first word in a sequence is either 'to' or '2', and if the second word in the sequence is tagged by the Twitter tagger as a verb (_V), and if that verb does not end in –*ing* (e.g. Look forward *to seeing* you). The tagger attaches the tag to the base tags assigned by the MDATT in the second step using another underscore and a tilda '~'. This tagger's precision rate is 95%, which was measured by computing the overall average of the accuracy of 100 random instances of each tagged feature (the lowest precision for a single feature was 67%). Examples (1iv)-(4iv) represent the output of the MDATT (the tag _~[TAG]) following the Gimpel et al. (2011) Twitter tagger (the first tag _[tag]).

1iv. @username_@_~INITIAL-MENTION
Sit_V_~VERB_~IMPERATIVE down_T_~MULTI-WORD-VERB
honey_N_~_NOUN_~GENERAL-NOUN ._,_~FULL-STOP

2iv. @username_@_~INITIAL-MENTION
Grow_V_~VERB_~IMPERATIVE up_T_~MULTI-WORD-VERB
lady_N_~NOUN_~GENERAL-NOUN ._,_~FULL-STOP

3iv. @username_@_~INITIAL-MENTION
This_O_~PRO_~DEMONSTRATIVE-PRONOUN

is_V_~VERB_~BE-MAIN-VERB

gold_A_~ADJ_~PREDICATIVE-ADJECTIVE [thumbs up emoji]_E_~EMOJI

4iv. @username_@_~INITIAL-MENTION

that's_L_~DEMONSTRATIVE-PRONOUN_~CONTRACTION_~BE-MAIN-VERB

not_R_~ADVERB_~NEG_~ANALYTIC-NEGATION

accurate_A_~ADJ_~PREDICATIVE-ADJECTIVE

This process for tagging was applied to both corpora of tweets. Having tagged the data, the next step in standard MDA is to measure and record the relative frequencies of each linguistic feature in each text. It is here that the second problem with applying MDA to tweets lies.

Table 3: The feature set of MDATT and the functional associations (based on Biber (1988) and CMC research)

Category	Feature	Description	Example	Functional associations
Tense and aspect markers	Past tense	Refers to verbs in their past tense form that are not in perfect aspect	went, saved, held	Associated with narratives (Biber, 1988).
	Perfect aspect	Refers to any form of HAVE + verb in past participle form	She *had been* to the shops already.	Describe actions completed in the past that are relevant (Quirk et al. 1985).
	Progressive	Refers to any form of BE plus (up to 2/3 adverbs and) verb ending in -ING	I *am talking* to Susan.	Describes ongoing action and is associated with spoken language (Collins, 2008).
Pronouns	1st person pronouns	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the first person: singular and plural plus contracted forms	I, We, us, me, myself, ourselves, ours, our, my, mine	Involved style and interpersonal focus (Chafe, 1982; Wales, 2006).

	2nd person pronouns	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the second person: singular and plural plus contracted forms	you, yours, you're, your	
	3rd person personal pronouns	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the third person: singular and plural plus contracted forms	he, she, theirs, themselves, them, hers	Anaphoric and deictic function (Wales, 2006). Associated with narrativity (Biber, 1988).
	Pronoun ‘it’	Refers to any form of pronoun IT: contracted, reflexive, possessive and possessive determiner	It is great, it's okay, itself, its	A non-personal gender (Quirk et al., 1985: 6.8).
	Demonstrative	Refers to the use of this, that, these, those as a pronoun; that is NOT followed by noun	*That* is my cat. I like it like *that*	Have definite meaning and thus assumes a shared context (Quirk et al., 1985).

	Indefinite pronouns	Refers to pronouns which indicate quantity or are indefinite pronouns	anything, somebody, I had a *few,* she had *several, some* of the men, *all* of the men	Lack definiteness and are quantitative (Quirk et al., 1985).
	Reflexive pronouns	Refers to pronouns in their reflexive form	myself, herself, ourselves, themselves, yourself	Anaphoric reference but can be used for emphatic purposes (Wales, 2006).
	WH pronoun			Can be used to form interrogatives and relative clauses. They are used for expansion (Chafe, 1985).
	Subject pronoun Nominative case	Refers to pronouns in their subject form	I, she, he, they, we	The agent or subject of the action and sometimes subject complement (Quirk et al., 1985)
	Object pronoun Accusative case	Refers to use of pronouns in their objective form	me, us, them, him	The object or patient acted upon. Can be associated with an informal style (Quirk et al., 1985).

	Possessive pronoun	Refers to pronouns which indicate possession	It is ours/mine/yours/theirs/his/hers.	Used to show possessive relation or ownership (Wales, 2006).
Questions	WH words	Refers to use of WH words	when, why, who, what, how	Interpersonal focus (Biber, 1986).
	WH-word + BE	Refers to WH-word + BE	Why are you going?	
DO	Question DO (WH-word + DO)	Refers to WH-word + DO	When do you care?	Questions in general have an involved style (Biber, 1986).
	Auxiliary DO	Refers to any form of DO that is followed by (up to three adverbs and) a verb.	I *do* not like cheese, I *did* take the bins out	Commonly occurs in the negative form or is used for emphatic purposes (Ard, 1982).
	Initial DO	Refers to when any form of DO is the first verb in the Tweet (after initial mentioning) or if DO is the first verb after a full stop	Doesn't the world look different	Associated with yes/no answers.

	Pro-verb DO	Refers to DO used as a main verb	I hate what he has *done*, She *did* it!	Associated with avoiding repetition and maintaining cohesion (Halliday and Hasan, 1976)
Nominal forms	Nominalisations	Refers to when verbs/adjectives are converted into nouns	action, statement	Associated with a high (abstract) informational focus (Biber, 1988).
	Numeral Noun	Refers to use of numerals functioning as nouns	I have three.	
	Ordinal Noun	Refers to use of ordinals functioning as nouns	I came first!	
	Nouns	Refers to other nouns that are not tagged as numeral, quantifiers, nominalisations, ordinals.		
	Proper nouns	Refers to anything tagged as a proper noun	Hillary_^ Clinton_^, London_^	
	Acronyms	Refers to any initials separated by full stops	U.S.A, U.S, N.Y.	

Place and Time adverbials	Place adverbials	Refers to adverbs indicating place	behind, beneath, downhill	Situates the content of the sentence in time and space (Biber, 1986). Modifies sentences or words (Virtanen, 1992). Time adverbs are used to provide clear and specific temporal information in situational contexts where time is relevant and when the audience is not physically present so as to communicate effectively (Bohmann, 2017).
	Time adverbials	Refers to adverbs indicating time	I'll be back *soon*	
Adjectives and Adverbs	Predicative adjectives	Refers to adjectives which come after a copular verb	I am *great*! She looks	Used for marking stance (Biber, 1988).

		crazy, Gyms smell *nasty*.	
Attributive adjectives	Adjectives that come before the noun and any other adjective not tagged as predicative.	The *big* cat	Highly integrative (Biber, 1988)
Comparatives	Refers to adjectives in their comparative form	She is *better* today, Drake>RiRi, I went for something *more* substantial	Comparatives and superlatives are used in evaluations, in particular for intensification and graduation (Martin and White, 2005).
Superlatives	Refers to adjectives and nouns in superlative form	the best, the worst, she is funniest	
Amplifier	Refers to adverbs used to intensify the verb/adjective	very, absolutely, so	The verb’s force is amplified or made more extreme (Quirk et al., 1985).
Downtoner	Refers to adverbs used to reduce the force of the adjective/verb	He is *slightly* fat, It was *pretty* awful	The verb’s force is lowered (Quirk et al., 1985).

	Usuality	Refers to adverbs indicating frequency and how often	always, never, often	To say how often something happens, specifying exact or indefinite time frame. Also used to express probability judgements (Cohen, 1999).
	Quantifying Adverb	Refers to quantifiers which are functioning as adverbs	You *all* are my inspiration, We are *all* happy to see you	Used to mark frequency or relative size.
	Other Adverb	Refers to other adverbs that are not tagged as amplifiers, downtoners, time and place adverbials, quantifying adverbs, adverbs of usuality.		Generally used for modification purposes (Biber, 1988).
Modals	Possibility modals	Refers to modals indicating probability/possibility/ability	can, may, mightn't	Used to mark possibility, ability or permission (Quirk et al., 1985).
	Necessity modals	Refers to modals indicating necessity/obligation	should, mustn't, ought	Marks necessity or obligation (Quirk et al., 1985)

	Predictive modals	Refers to modals indicating prediction	will, shall, I'll	Marks volition or prediction (Quirk et al., 1985)
Specialised Verb Classes	Public verbs	Refers to public verbs: used to report on speech	told, said, shouted	Introduce indirect statements (Biber, 1988).
	Private verbs	Refers to private verbs: used to encode feelings, opinions, emotions, cognition	believe, think, know, learn	Express intellectual states or non-observable intellectual acts (Biber, 1988).
	Suasive verbs	Refers to verbs which refer to persuasion	beg, insist, command, demand, allow…	Used to bring about some future change (Biber, 1988).
	Phrasal verbs	Refers to both prepositional and particle verbs	Screwed up, keep up, hang out.	Idiomatic properties.
	Verbs of perception	Refers to verbs of perception	hear, smell, taste	Used to encode experience (Viberg, 2009). Feel, see, make, watch, hear, help

	Stance verbs	Refers to verbs used to encode stance	want, seem, appear, like, love, prefer, need …	Used to mark stance (Biber, 2006). Tend, happen, seem, appear, want
Modifiers	Indefinite article	Refers to use of indefinite article	a, an	Used to determine nouns (Quirk et al., 1985).
	Definite article	Refers to the use of the definite article	*The* cat	
	Possessive determiner	Refers to determiners which indicate possession	our cat, your house, their garden, his eyebrows	Used to indicate possession of nominal referent (Quirk et al., 1985)
	Quantifiers	Refers to quantifiers used as a determiner	*Few* people, *some* people	Can have emphatic properties (Biber, 1988).
	Prepositions	Refers to the use of prepositions	*down* the road, *in* your car	Packing in high amounts of information (Biber 1988). Used for noun modification.
	Titles	Refers to titles	Mr. Dr, Miss, Sir	Used to modify and signal rank or status (Quirk et al. 1985).

	Numeral determiners	Refers to use of numerals functioning as determiners	Three dogs	Function as either heads in a noun phrase or as determiners (Quirk et al., 1985).
	Ordinal determiners	Refers to use of ordinals functioning as determiners	She took second place.	Can signal rank or date.
	Pre-determiners	Refers to determiners which come before determiners	*All* the people in this room are intelligent.	Can have emphatic purposes.
	Quantifying pre-determiners	Refers to quantifier as a pre-determiner	*All* the people in this room are intelligent.	Used to mark quantity or relative frequency of the nominal referent.
	Demonstrative Determiner	Refers to *this, that, these, those* followed by a noun (which can be preceded by adjectives, adverbs).	*That* smelly cat	Have definite meaning and thus assumes a shared context (Quirk et al., 1985).
Coordination	Contrastive conjunctions	Refers to conjunctions that signal a contrast is being made	but, by contrast	Used to show a contrast or difference.
	Coordinating conjunctions	Refers to coordinating conjunctions	and, and	Used to mark logical relations between clauses.

	Other conjunctions	Refers to other conjunctions not tagged as either contrastive or coordinating	In addition to, e.g., thus	Used to mark logical relations between clauses and can have a highly informational focus (Biber, 1988)
Subordination	Cause subordinators	Refers to subordinators which indicate a causal relationship	Because, ‘cause	To indicate a cause or reason.
	Time subordinators	Refers to subordinators indicating time	*While* his mother slept, he snuck out the window.	Used to express time. Can be common in procedural texts (Quirk et al., 1985)
	Place subordinators	Refers to subordinators indicating place	I will find you *wherever* you go, you can find me *where* the food table is.	Used to indicate position or direction (Quirk et al., 1985).
	Conditionals	Refers to subordinators indicating a condition	if, unless	Introduces a possibility (Finegan, 1982).
	Concessive	Refers to subordinators which indicate concession	although, though	Can be used for framing purposes and for introducing background information (Biber, 1988).

Possession	Possessive pronoun	Refers to pronouns which indicate possession	It is ours/mine/yours/theirs/his/hers.	Mark possession. Involved and interpersonal function (Biber, 1988).
	Possessive proper noun	Refers to proper nouns in possessive form	*Donald Trump's* hair	
	Possessive determiner	Refers to determiners which indicate possession	our cat, your house, their garden, his eyebrows	
	Possessive noun	Refers to nouns in possessive form	The cat's dinner	
	Possessive quantifying pronoun	Refers to quantifying pronouns in their possessive form	somebody's jumper	
Relatives	Relative clause subject gap	Refers to relative clauses with subject gap	The man that was cursed	Allow for more exact and explicit reference (Ochs, 1979), as well as idea unit expansion and integration (Chafe 1982; 1985).
	Relative clause object gaps	Refers to relative clause with object gap	The man that the gypsy cursed	

	Pied-piping relative	Refers to the use of preposition + relative pronoun to avoid stranded preposition	with/to whom did Sarah speak?, the box in which it was kept.	
Other Verbs	Have as main verb	Refers to when any form of HAVE is the main verb	She *has* so much money, I *had* seven chocolates	Used to signal a relationship of possession (Butt et al., 2003).
	Be as initial verb	Refers to when BE is the first verb in the Tweet (after initial mentioning) or if BE is the first verb after a full stop.	Am going to the shops. Is it okay?	Can indicate pronoun omission or can mark a question. Omission of subject pronouns and sometimes auxiliaries are associated with an informal spoken style and thus may be employed as a way to reflect orality and phonological reduction (Werry, 1996).
	Have as initial verb	Refers to when any form of HAVE is the first verb in the Tweet (after initial mentioning) or if HAVE is the first verb after a full stop	*Had* an absolute nightmare! *Have* you been to the shops?	
	Initial verb-ing	Refers to when verb ending in -ing is the first verb in the Tweet (after initial mentioning) or if it is the first verb after a full stop	Going out.	

Initial verb	Refers to initial verbs in their base form which are followed by particular things making them unlikely to be imperative clauses	wish you were here, love to go, want to spend, do you
Initial verb question	Refers to when particular auxiliary and dummy auxiliary verbs are the first verb in tweets (after initial mentioning) or if it is the first verb after a full stop	ain't, aren't, aint, is, arent, am, are, had, hadn't, hadnt, has, hasnt, hasn't, isn't, did, didn't, didnt, haven't, havnt, havent, doesn't, doesnt
Initial verb Third person singular	Refers to when verb ending in -s is the first verb in the Tweet (after initial mentioning) or if it is the first verb after a full stop (except for a select few verbs which can be used for imperatives)	@username thinks she is fab. Takes less time to watch paint dry.

	Initial verb past tense	Refers to when verb in past tense/past participle form is the first verb in the Tweet (after initial mentioning) or if it is the first verb after a full stop	Wanted that for ages. Said like a true gentleman.	
	Initial modal verb	Refers to when modal verb is the first verb in the Tweet (after initial mentioning) or if a modal verb is the first verb after a full stop	Could you scratch my back?	Can be used to make a request.
	Verb-ing	Refers to verb in ING form that is not in standard progressive form (likely a gerund/nominalisation or auxiliary and pronoun omission)	Going for walks is my favourite thing to do on a Saturday	Can indicate auxiliary and pronoun omission but can also indicate gerund.
	BE going to	Refers to any form of BE (including contracted) + going + to	I*'m going to be* in Kansas tonight, She *is going to* leave her job.	Used to encode future intentions and mark prediction.
	Third person singular verb	Refers to verbs ending in -s	thinks, has, takes	Marks present tense. Used to deal with topics of immediate relevance (Biber, 1988)

Stative forms	Be as main verb	Refers to when BE is the main verb and when BE is in its copular form; that is, when it is followed by a predicative adjective	She *is* a beautiful woman; She *is* beautiful	Used to introduce entities or describe their characteristics or attributes (Butt et al., 2003). Used to encode predicative descriptions of a subject. Associated with a fragmented and unplanned production of text (Biber, 1988).
	Existential ‘there’	Refers to the use of *there* in its existential form and thus not as a place adverb	*There* was a man in dark clothing, *There* may be 5 or 6 obstacles	
	Copular verbs	Refers to copular verbs but not BE as a main verb (even if it is in its copula form, that is: when it is followed by predicative	*Appear, seem, taste, grow, keep, got*	
Mood	Imperatives	Refers to clauses in imperative mood	Go away!, Don't be foolish! Take care.	Spertus (1997) found that imperative statements tend to be insulting. To make demands, associated with procedural texts (Butt et al., 2003). Can indicate pronoun omission.
Negation	Synthetic negation	Refers to use of nor, neither and no - but not as interjection	No, neither, nor, no more	

	Analytic negation	Refers to 'not' plus contracted forms	can't, cannot, not	Synthetic is more integrated, whereas analytic is more fragmented (Tottie, 1983).
Interjections	Positive interjections	Refers to any form of YES tagged as an interjection by the Gimpel tagger	Yeahhhh_!, Yup_!, Ya_!	Fillers used to gain time, maintain conversation and/or to show attentiveness (Smith, 2003). Some types of interjections, for example ‘hmmm’ has been shown to convey scepticism with the function of withholding agreement and thus mitigating disagreement (Vandergriff, 2013). Turn-initial ‘no’ can be used to show disagreement or rejection (e.g. No, you’re wrong), to show that someone has not necessarily understood the original message (e.g. ‘No, I don’t mean X, I mean Y’) (Schegloff, 1992), to mark moving from a joke to a serious tone (e.g. ‘No but on a serious note’) (Schegloff, 2001), to acknowledge someone else’s talk or affiliate with it (Jefferson, 2002).
	Negative interjections	Refers to forms of NO that are tagged as interjections by Gimpel tagger	No_!, Naaaa_!	
	Other interjections	Refers to other interjections that are not tagged as laughter, positive interjection 'Yes', negative interjections 'No'	OMG, WOW!	

	Laughter			Laughter occurs in assessment environments (Petitjean and Morel, 2017), to show understanding or appreciation of a joke (Norrick, 1994), to highlight irony (Uygur-Distexhe, 2012), a phatic filler (Baron 2004), to show disaffiliation (Holt, 2012), a positive assessment or affiliation (Petitjean and Morel, 2017), to orient to the previous message as laughable (Petitjean and Morel, 2017), to show that the previous message is being taken as non-serious (Petitjean and Morel, 2017), it can be a resource for turn-taking (Petitjean and Morel, 2017).
		Refers to written out laughter	haha, lol, lmao, lmfao	
Punctuation	Question marks	Refers to the use of question mark	?, !?, !?!	Reflects paralinguistic cues in CMC (Smith, 2003).

	Exclamation marks	Refers to the use of exclamation marks	!!!, !?	Multifunctional (e.g. Waseleski, 2006; Vandergriff, 2013): to express surprise (Smith, 2003), excitability, friendly closings, aggravated disagreement, and cues to humour (Vandergriff, 2013).
	Quotation	Refers to the use of quotation marks (single/double)	‘’ “”	Can be used to refer to direct speech, or they can be used for ironic effect.
	Capitalisation	Refers to two or more capital letters that is not tagged as an acronym/ URL/ mentioned username	HAPPY	Used for emphasis (Smith, 2003) or to denote shouting (Postmes et al., 2000).
	Colon	Refers to the use of colons	:	Used to introduce a list, definition, description or explanation. Can also be used on social media to introduce extra content (e.g. URL).

	Semi-colon	Refers to use of semicolon	;	Used to introduce close relation between two independent clauses. Used to introduce a list.
	Comma	Refers to the use of commas	,	Used to introduce a clause or main sentence. Associated with informational elaboration.
	Brackets	Refers to the use of brackets	()	Used for supplementary information for the purpose of clarification or exemplification.
	Ellipsis	Refers to three or more full stops	…	Used to omit part of sentence. Used to create suspense. Used to indicate a brief pause.
	Full stop	Refers to use of full stop	.	Used to indicate sentence ending/boundary. Can suggest multiple sentences.
CMC	URLs	Refers to URLs: can be meme, gif, status, link to website, video etc.		To expand Tweets (Yazdanfar and Thomo, 2013). URLs are employed to recommend articles in real-time (Sankaranarayanan et al., 2009).

	Emojis/Emoticons	Refers to anything tagged by the Gimpel tagger as an emoticons and some unicodes.	:)	Deliberately used (Dresner and Herring, 2010). Communicate humour of solidarity or display sarcasm (Wolf, 2000), to emphasise or clarify a particular emotional state, to soften a negative tone, or to regulate the interaction (Derks et al., 2008). Used to promote politeness (Darics, 2010), to mark affect or to orient to dispreferred action (Vandergriff, 2013).
	Hashtags	Refers to the use of Hashtag	#hashtag	Hashtags are used to annotate Tweets to specify the topic or intended audience of the message (Conover et al., 2011). Hashtags are linked to a stream of content and thus users contribute to and participate in the stream when they choose to use one (Conover et al., 2011).
	Mentioning: initial and non-initial	Refers to Tweet initial mentioning	\@username how are you?	To directly address another user as well as (although rare) refer to an individual in the third person (Honeycutt and Herring, 2009).
		Refers to mentioning that is not initial	Is \@username even here	

	Verb-initial	Refers to any verb in initial position of a tweet or after initial mentioning.	Thought about it…no. Am in London.	Omission of subject pronouns and auxiliaries are associated with an informal spoken style and thus may be seen to be employed as a way to reflect orality and phonological reduction (Werry, 1996).
Complementation	That verb complements	Refers to Private verbs, Public verbs, or suasive verbs + that	I *think that* you are pathetic for sleeping with a night light.	Used to expand an idea-unit (Chafe, 1982; 1985). Informational elaboration (Biber, 1988). Can serve interpersonal functions (Biber, 1986).
	That adjective complements	Refers to adjective that complement clauses	It's *pathetic that* you can't sleep without a night light at 40.	
	Noun+ that complements	Refers to noun complement clauses	The *fact that* you can't sleep without a night light makes you pathetic.	
	*Adjective + to*complements	Refers to adjective + to complement clause	I am happy to go with Karen.	

	Infinitives	Refers to verbs in infinitive form that is not adjective + to complement clause or split infinitive	to be, to have	
	Split infinitives	Refers to verb in infinitive form separated by adverb(s)	to really hate, to not like	
	WH-clauses	Refers to WH clauses	Do you understand *what cooperation is?*	
	Gerund	Refers to prepositional complement: when a preposition is followed by noun in -ing form (but this is tagged by Gimpel tagger as a verb)	Sarah talked about *leaving* her job	Used as prepositional complement. Informational elaboration and interpersonal function (Biber, 1988).
Passive constructions	Agentless passives	Refers to use of passive voice without the inclusion of an agent	He was arrested. She was told not to speak.	Associated with a detached style (Biber, 1988). Agent is either given prominence or is removed

	By passives	Refers to use of passive voice with agent in by clause	He was arrested by the police. She was told not to speak by her teacher.	from the sentence (Fairclough, 1992).
Contractions	Pronoun with verb contracted	Refers to when the verb is contracted with pronoun	I'm, She'd, They've, You'll, That's	Reduced surface form (Biber, 1986). From a prescriptivist perspective, they are dispreferred in certain registers (e.g. academic writing) (Finegan, 1980). They have been found to occur more frequently with interactive features (Biber, 1988).
	WH- word with verb contracted	Refers to WH that have the verb contracted	*what's, who'd, where's*	
	Quantifying pronoun with verb contracted	Refers to quantifying pronouns with the verb contracted	noone's happy today, everyone's been before	
Profanity	Potential swear words.	Refers to words that can be used to offend/abuse as well as swear words generally. They may also be used harmlessly	Fuck, cunt, twat	Profanity can be used: to abuse, for emphasis, for reclamation, mark exasperation and excitedness (Clarke and Grieve, 2017; Waseem et al., 2018; Ajayi, 2018).

4.6. Problem 2: Analysis of Short Texts

The second problem with applying the standard version of MDA to individual tweets is that tweets are exceptionally short texts, rarely exceeding 40 words. MDA is based on the normalised frequency counts of all linguistic features to a text length of 1,000 words because texts can vary considerably in length. Non-normalised counts do not represent comparable frequencies of occurrence, but instead only provide raw frequencies. For example, text A is 1,000 words long and text B is 3,000 words long, and both texts contain 50 nouns. Despite both having the same raw frequency, nouns do not occur at the same rate in these texts, rather they occur three times as frequently in text A than text B. It is therefore important to compute the relative frequencies of features before running the factor analysis in standard MDA, as otherwise the results will be skewed. However, the relative frequencies of linguistic features tend to only be meaningful when the text is over a certain length in words, as the relative frequencies of features in shorter texts are not reliable estimates of their relative frequencies in a larger sample of similar texts more generally.

For example, the average length of the tweets within the corpus of general tweets is 17 words long. In a tweet containing 17 words, each word will have a relative frequency of at least once per 17 words. However, even the most frequent words in a larger corpus of tweets will come nowhere near this rate. Moreover, every word that does not occur in that tweet will have a relative frequency of 0. Many of those words, however, might occur much more frequently in a larger sample of tweets. Essentially, the relative frequencies of most features cannot be measured in short texts because they are too rare, which means that we do not see enough tokens in any given text to get a reliable measure of relative frequency. As a result, most studies employing MDA have tended to be limited to analysing longer texts, such as texts of around 500 (Passonneau et al., 2014) or 1,000 words long (Biber, 1993).

Nevertheless, there have been some studies, which have examined short texts using MDA, including tweets (e.g. Passonneau et al., 2014; Friginal et al., 2018; Coats, 2016; Titak and Roberson, 2013). These studies, however, have combined the short texts to form longer text chunks that are ultimately more suitable for frequency-based analyses. This approach is valid if texts are combined in a principled manner. However, the manner of concatenation in these studies is not always made clear and often seems that they are combined just to make up text samples of the required length for frequency-based analyses.

Concatenation tends to limit the kinds of research questions to comparing tweets generically or groups of tweets to other registers and varieties, such as online and offline language, as opposed to investigating variation between individual tweets. For example, Coats (2016) combined tweets to create 1,000 word chunks and compared tweets from Finnish English authors with English tweets from authors from across the world. In another study, Passonneau et al. (2014) compared tweets to other registers. They indicate that they initially combined tweets to make up 1,000 words chunks following Biber's (1993) suggestion that even rare forms are relatively stable at this length, however they found that 500 word chunks produced stable frequency rates for their feature set.

Sardinha (2014) combined three tweets and labelled this a Twitter text unit in his study examining various Internet registers along the dimensions of variation in Biber (1988). This was based on the fact that, at the time of data collection, the total users (140 million) produced 340 million tweets, which is an average of 2.4 tweets per user per day, which they rounded up to 3. Despite this rationale about the amount of tweets per user per day, the Twitter text units in Sardinha's (2014) study are not actually made up of 3 tweets from the same author, but instead can be from three different people. A tweet's communicative function can vary from one tweet to the next and a single author may employ a different style to the next. Overall, whilst concatenation can still be informative when done in a principled manner, it largely limits the kinds of descriptions that can be made. Specifically, concatenating tweets does not enable the identification and description of functional linguistic variation across individual tweets, which this dissertation seeks to examine.

The issue of relative frequencies in short texts is not new. It is an important methodological issue in stylometry (Stamatatos, 2009), such as in authorship analysis (Grieve et al., 2018a), authorship attribution (Schwartz et al., 2013; Layton et al., 2010), authorship verification (Brocardo et al., 2013) and forensic linguistic analysis (Coulthard et al., 2017; Ehrhardt, 2007). Different solutions and methods have been offered, including looking at idiosyncratic and consistently used features (Grant, 2013), shared word-sequences (Nini, 2018), and the presence or absence of particular linguistic features (e.g. Layton et al., 2010; Schwartz et al., 2013; Brocardo et al. 2013) in particular random samples of texts from individual author corpora (Grieve et al., 2018a), as opposed to their relative frequencies.

4.7. The Solution: Short-Text Multi-Dimensional Analysis

In order to identify the variation between tweets, this dissertation introduces a new form of MDA that allows for the variation between individual short texts to be identified and described (see also Clarke and Grieve, 2017; Clarke, 2018; Clarke, 2019). Specifically, rather than measure the relative frequencies of features, the method examines whether the feature is present or absent and records this information in a categorical data matrix. Subsequently, rather than use factor analysis on this matrix, which is used for continuous data (e.g. relative frequencies of features), this approach uses Multiple Correspondence Analysis (MCA), which is similar to factor analysis, but unlike factor analysis[1], MCA works specifically with categorical data (e.g. presence/absence of features). MCA is used in this dissertation like factor analysis in standard MDA – to reveal the major sets of co-occurring linguistic features and the texts most strongly associated to these patterns, which, like in standard MDA, are then interpreted for the underlying communicative function by analysing the particular linguistic co-occurrence patterns in the context of the tweets associated to the patterns (described in more detail below).

Specifically, for each corpus of tweets, a separate data matrix was created, recording the presence and absence of 124 linguistic features tagged by the MDATT. Le Roux and Rouanet (2010) advise that very infrequent

[1] Different techniques have been offered so that factor analysis can be applied to categorical data (see Bartholomew, 1980; Mislevy, 1986).

features (e.g. those that occur in < 5% of the data) either need to be pooled with other related features or they might need to be discarded because infrequent features can overly influence the results of MCA. Thus, the linguistic features that occurred in fewer than 5% of the tweets were either pooled or removed. The decisions for pooling and removing features can be found in Appendix 1. After this pooling process was completed and the final feature set was decided on, each tweet was analysed for the presence or absence of the linguistic features occurring in more than 5% of the tweets (see Appendix 1 for final feature sets) using another computer programme written in Perl developed to record this information in a data matrix. The remaining linguistic feature by tweet data matrix for each corpus was then subjected to MCA.

4.7.1. Multiple Correspondence Analysis (MCA)

MCA is a geometric data analytic method, which enables the identification and visualisation of the most dominant relationships between three or more categorical variables in a low-dimensional space. The method was popularised by Jean-Paul Benzécri, who used it to analyse sociological data from questionnaires (Benzécri, 1979), as it can be used to observe relationships between individuals (e.g. people who have answered similarly or dissimilarly to the questions), as well as to visualise the relationships between the variables (i.e. which answers to the questions tend to be selected together, and which answers are rarely selected together). Specifically, Benzécri used (1979) MCA to visualise the relationships between people and their responses to questions in terms of distance, producing two clouds of points, where the points on one cloud represent the people, and the points on the other are the responses to the questions, and the distance between each point is based on how similar they are to each other in their distribution. Points representing people are closer together in the space if they give the same responses to the questions. Points representing responses to questions are closer together if they are distributed similarly across the people. In other words, if many people select the same responses then those responses are closer together in the space.

In addition to analysing data from surveys or questionnaires (Greenacre and Pardo, 2006), MCA has been used in a range of exploratory studies, including those concerned with the identification of factors contributing to motorcycle crashes (Jalayer and Zhou, 2017), different tastes (Le Roux and Rouanet, 2010; Le Roux et al., 2008), different patterns of cultural consumption (Kahma and Toikka, 2012), patterns of ageing (e.g. Costa et al., 2013; Sourial et al., 2010), and for linking crimes (e.g. Yokota et al., 2016). Moreover, MCA has been used in a small number of linguistic studies, mainly to identify confounding variables (Tummers, Speelman and Geeraerts, 2012), and to identify patterns of usage of polysemic words (Glynn, 2009).

In general, MCA identifies the most dominant patterns of variation within a data matrix of individuals *I* and categorical variables *V* (Le Roux and Rouanet, 2010). In this dissertation, the individuals *I* are the set of *n* tweets and the categorical variables *V* are the linguistic features occurring in more than 5% of those tweets. Each linguistic feature has two categories *k* and *k*`, namely *presence* and *absence*. Hence, a cell in the matrix (*i, v*) reflects whether the linguistic feature *v* is present or absent in the tweet *i*. Based on the data matrix, MCA produces two clouds of points: a cloud of tweets and a cloud of categories of linguistic features, which positions each tweet and each category of a linguistic feature geometrically as a point in a low-dimensional space. The rest of this section describes the ways in which distance is worked out in MCA for these two clouds, as well as the dimensions, the coordinates and contributions, and supplementary elements, which are used in the interpretation of the results. The application of the method in R is demonstrated on the four example trolling tweets from section 4.5.

<u>The cloud of tweets (Le Roux and Rouanet, 2010)</u>

The distance *d* between tweets in the cloud of tweets is defined as follows. If for linguistic feature *v,* tweet *i* and *i*` both have the presence of *v,* the part of the distance due to linguistic feature *v* is null: $d_v(i, i`) = 0$. The distance between two tweets will be based on the linguistic features that they do not share. In other words, when tweet *i* has the presence of linguistic feature *v,* denoted by *k,* and tweet *i*` has the absence of it, denoted by *k`,* then the two tweets will be positioned further apart. Specifically, in this situation where they do not share the category linguistic feature *v*, the part of the squared distance between tweet *i* and *i*` due to linguistic feature *v* is defined by the formula

Equation 1 $$\boldsymbol{d_v^2(i, i`)} = \frac{\mathbf{1}}{\boldsymbol{fk}} + \frac{\mathbf{1}}{\boldsymbol{fk`}}$$

where $fk = n_k/n$ is the relative frequency of tweets that have the presence of linguistic feature *v,* denoted by *k,* and $fk` = n_{k`}/n$, which is the relative frequency of tweets that have the absence of linguistic feature *v,* denoted by *k`*.

The overall squared distance between tweet *i* and *i`* takes into account all linguistic features, denoted by *V*. In other words, the overall squared distance between tweet *i* and *i`* will be based on the categories of all the linguistic features in the data matrix that they do not share. This squared distance is defined by the formula

Equation 2 $$\boldsymbol{d^2(i, i`)} = \frac{\mathbf{1}}{\boldsymbol{V}} \sum_{\boldsymbol{v} \in \boldsymbol{V}} \boldsymbol{d_v^2(i, i`)}$$

The set of all distances between tweets determines the cloud of tweets consisting of *n* points (number of tweets).

If G is the mean point of the cloud and M^i denotes the point representing tweet *i,* the squared distance from point G to M^i is

Equation 3 $$(\mathbf{G}\boldsymbol{M^i})^2 = \left(\frac{\mathbf{1}}{\boldsymbol{V}} \sum_{\boldsymbol{k} \in \boldsymbol{K_i}} \frac{\mathbf{1}}{\boldsymbol{fk}}\right) - \mathbf{1}$$

where K_i denotes which linguistic features out of all the linguistic features are present or absent in tweet *i* (i.e. the set of *V* categories in tweet *i*).

The variance of the cloud, also termed the eigenvalue, is the mean of the squared distances from the points of the cloud to the mean point, defined by the formula

Equation 4 $$\sum(\mathbf{G}\boldsymbol{M}^{\boldsymbol{i}})^{\mathbf{2}}/\boldsymbol{n}$$

A basic characteristic of a cloud of points is its dimensionality (Le Roux and Rouanet, 2010). The dimensionality of the cloud is *L*, where $L \leq K - V$ (overall number of *K* categories minus the number of variables *V*). For instance, a data set of 63 linguistic features each with two categories, namely presence and absence, could yield a cloud with less than or equal to 63 dimensions. Importantly, the dimensionality of the cloud of tweets and the cloud of categories of linguistic features are the same in MCA.

<u>The cloud of categories of linguistic features (Le Roux and Rouanet, 2010)</u>

The cloud of categories of linguistic features is a weighted cloud of *K* points. Category *k (*e.g. presence of nominalisation) is represented by a point denoted by M^k with weight n_k, which denotes the amount of tweets that have category *k*. For each linguistic feature, the sum of the weights of category points is *n* (i.e. the total number of tweets)*.* The whole set of categories *K* is the sum $n\mathrm{V}$*.* The relative weight p_k of point M^k is

Equation 5 $$\boldsymbol{p_k} = \frac{\boldsymbol{n_k}}{\boldsymbol{n}\mathbf{V}} = \frac{\boldsymbol{fk}}{\boldsymbol{V}}$$

The squared distance between two points on the cloud of categories of linguistic features M^k and $\mathrm{M}^{k`}$ is denoted by the formula

Equation 6 $$\left(\mathbf{M}^{k}\mathbf{M}^{k`}\right)^{2} = \frac{(n_k + n_{k`} - 2n_{kk`})}{n_k n_{k`}/n}$$

where $n_{kk`}$ refers to the number of tweets that have both of the categories of linguistic features. When *k* is the presence of linguistic feature *v* and *k`* is the absence of the same linguistic feature then $n_{kk`} = 0$ and the squared distance between the points is

Equation 7 $$\left(\mathbf{M}^{k}\mathbf{M}^{k`}\right)^{2} = \left(\frac{1}{fk}\right) + \left(\frac{1}{fk`}\right)$$

The mean point of the cloud of categories is also denoted by G. The squared distance from M^{k} to G is

Equation 8 $$\left(\mathbf{GM}^{k}\right)^{2} = \frac{1}{fk} - 1$$

The cloud of categories and the cloud of individuals have the same variance. The variance of the cloud is

Equation 9 $$\sum p_k \left(\mathbf{GM}^{k}\right)^{2} = \left(\frac{K}{V}\right) - 1$$

Thus, for example, the variance of the clouds in a data set with 63 linguistic features (126 categories) is equal to $\left(\frac{126}{63}\right) - 1 = 1$.

<u>Principal Axes (Le Roux and Rouanet, 2010)</u>

Clouds of points can be projected onto principal axes. Whilst the midpoints between projected points are preserved from the cloud of points, fitting the data onto an axis loses information from where the point should be in a multidimensional cloud, and this is called residual deviation (Le Roux and

Rouanet, 2010). Each cloud's principal axes are ranked in decreasing order of importance (i.e. the best fit of the data), so that principal axis 1 is the most important and best one-dimensional fit of the data, where the sum of the squared residual deviations of points is at the minimum (Le Roux and Rouanet, 2010).

Each dimension of a cloud, denoted by l, explains a proportion of the variance in the data and this is denoted by the eigenvalue λ_l. The total variance of the cloud is the sum of the eigenvalues $\sum \lambda_l = \frac{K}{V} - 1$. The mean eigenvalue is $\bar{\lambda} = \frac{\frac{K}{Q} - 1}{K - Q} = \frac{1}{Q}$. Because the clouds tend to have a high dimensionality, the variance rates of the dimensions are usually quite low (Le Roux and Rouanet, 2010). Consequently, Benzécri (1992: 412) introduced modified rates, which reveal more clearly the importance of the first few dimensions. Modified rates only consider the axes whose eigenvalues are above the mean. Specifically, for $l = 1, 2, \ldots l_{max}$ such that $\lambda_l > \bar{\lambda}$, calculate

Step 1: The pseudo-eigenvalue $\lambda'_l = \left(\frac{Q}{Q-1}\right)^2 \left(\lambda_l - \bar{\lambda}\right)^2$;

Step 2: The sum $S = \sum_{l=1}^{l_{max}} \lambda'_l$;

Then, for $l \leq l_{max}$, the modified rates are equal to $\tau'_l = \frac{\lambda'_l}{S}$

For each dimension, each tweet *i* and each category *k* of linguistic feature *v* is assigned a positive or negative coordinate. The coordinate of the tweet's point M^i relative to dimension l is denoted by y^i_l. The coordinate of the category of linguistic feature point M^k relative to dimension l is denoted by y^k_l.

For each dimension, the coordinate y^i of a particular tweet point M^i is the simple mean of the coordinates y^k of the *V* category points K^i in tweet *i* (i.e the set of *V* categories in tweet *i*), divided by the square root of the eigenvalue $\sqrt{\lambda}$:

Equation 10 $$\boldsymbol{y^i} = \frac{\mathbf{1}}{\sqrt{\boldsymbol{\lambda}}} \sum_{\boldsymbol{k} \in \boldsymbol{K_i}} \frac{\boldsymbol{y^k}}{\boldsymbol{V}}$$

For each dimension, the coordinate y^k of the category of linguistic feature point M^k is the simple mean of the coordinates y^i of the n_k points of tweets $(i \in I_k)$ that have that particular category of linguistic feature, divided by the square root of the eigenvalue $\sqrt{\lambda}$:

Equation 11 $$\boldsymbol{y^k} = \frac{\mathbf{1}}{\sqrt{\boldsymbol{\lambda}}} \sum_{\boldsymbol{i} \in \boldsymbol{I_k}} \boldsymbol{y^i}/\boldsymbol{n_k}$$

The distance between the coordinates of tweets on each dimension indicates the dissimilarities in their linguistic composition with respect to the major pattern of variation that the dimension represents. Essentially, the shorter the distance between the tweets in the space indicates that these tweets are more similar in distribution (Costa et al., 2013); that is, the tweets tend to have the same categories of linguistic features. With respect to the categories of linguistic features, the coordinates reflect the nature of the association between the categories of linguistic features in terms of proximity, where linguistic features that are distributed in similar ways in the tweets will have coordinates closer to each other (Le Roux and Rouanet, 1984; 1988; 2010).

In addition to coordinates, MCA assigns each tweet and each category of a linguistic feature a contribution for each dimension, which denotes the proportion of variance of the dimension due to that point (Le Roux and Rouanet, 2010); that is, how much of the variability of the dimension the point explains. If *p* denotes the relative weight of a point, and *y* its coordinate relative to the dimension of variance λ, the contribution of point to axis is equal to $(p\ y^2)/\lambda$. Specifically, the contributions are calculated by dividing the squared distance from the mean point to the particular point by the total amount of points, and then this is divided by the variance of the dimension, also termed the eigenvalue (Le Roux and Rouanet, 2010). Hence, the contribution of tweet *i* point M^i is

Equation 12 $$\mathbf{Ctr}_{\boldsymbol{i}} = \frac{\frac{1}{n}(y^i)^2}{\lambda},$$

and the contribution of category of linguistic feature point M^k is

Equation 13 $$\mathbf{Ctr}_{\boldsymbol{k}} = (\frac{fk}{V}\left(y^k\right)^2)/\boldsymbol{\lambda}.$$

Contributions show which categories of features and which tweets are the most important contributors to the dimensions. In this way, they are similar to factor loadings in factor analysis. The contributions of points to each dimension are the main aid to interpretation. Le Roux and Rouanet (2010) suggest that the categories of variables contributing above the average contribution should be interpreted, as these represent the most distinguishing patterns of variation. All the contributions for each dimension are positive numbers and equal 100. Therefore, the average contribution of the categories of features is $\frac{100}{K}$. Together, the contributions and coordinates of the tweets and categories of linguistic features returned by MCA reveal the range of the most important patterns of linguistic co-occurrence in the corpora of tweets, and the tweets most associated with these patterns. Like Biber’s MDA, the results of the MCA are then interpreted. Benzécri (1992: 405) noted:

> *Interpreting an axis amounts to finding out what is similar, on the one hand, between all the elements figuring on the right of the origin and, on the other other hand, between all that is written on the left; and expressing with conciseness and precision the contrast (or opposition) between the two extremes.*
>
> (Benzécri, 1992: 405)

Thus, because the distance between the features’ coordinates reflects their co-occurrence in the dataset, the features most strongly contributing to the dimension with positive coordinates are interpreted in opposition to the features with negative coordinates that have strong contributions. This is repeated for each subsequent dimension until the dimensions are no longer readily interpretable. Other studies employing MCA have also similarly interpreted the coordinates and contributions of categories variables in this way (e.g. Le Roux and Rouanet, 2010; Jalayer and Zhou, 2017; Kahma and Toikka, 2012).

MCA for MDA Demonstration

To illustrate, using ‘FactoMineR’ in R MCA was applied to the tagged examples (1iv)-(4iv) in section 4.5, which are reproduced here in their bare form for reference purposes.

1. @username Sit down honey.
2. @username Grow up lady.
3. @username This is gold

4. @username that’s not accurate

Table 4 represents the categorical data matrix of these four example tweets, where each tweet is represented in a row and each linguistic feature is denoted by a column, and each cell indicates whether that feature is present (P) or absent (A) in the tweets (1iv)-(4iv).

Table 4: The data matrix used in the MCA demonstration representing the occurrence of features in tweets (1iv)-(4iv) (P = present, A = absent)

	Analytic-Negation	**Be-Main-Verb**	**Demonstrative**	**Emoji**	**Full Stop**	**Imperative**	**Initial-Mention**	**Multi-word-verb**	**General-noun**	**Predicative-Adjective**	**Contraction**
1iv	A	A	A	A	P	P	P	P	P	A	A
2iv	A	A	A	A	P	P	P	P	P	A	A
3iv	A	P	P	P	A	A	P	A	A	P	A
4iv	P	P	P	A	A	A	P	A	A	P	P

Based on the data matrix, the MCA produced a cloud of tweets and a cloud of categories of linguistic features, whose dimensionality *L* is no more than 10 ($L \leq 21 - 11$). Table 5 presents the eigenvalues of the analysis, the percentage of variance explained by the dimensions, and the cumulative percentage of them. Eigenvalues provide a summary of the data matrix and their sum represents the variance of the cloud. The percentage of variation reveals how much of the variation in the data matrix each dimension captures. Table 5 shows that the four Tweets comprised of 21 categories of 11 variables can be explained in two dimensions with the first dimension being able to describe 80% of the variation found across the tweets.

Table 5: The eigenvalues and the percentage of variance explained by the dimensions in the demonstration of MCA for MDA

	Dim.1	**Dim.2**
Eigenvalue	0.731	0.178
%_of_variance	80.368	19.632
Cumulative_%_of_variance	80.368	100

Table 6 shows the results of the MCA on the categories of linguistic features for Dimension 1 on examples (1iv)-(4iv). Notably, for each dimension the MCA assigned each category of a variable (e.g. Emoji_P and Emoji_A) a positive or negative coordinate, as well as a value indicating its contribution to the dimension. The shorter the distance between the coordinates of the categories of linguistic features indicates that these categories of linguistic features are distributed in similar ways in the tweets (Le Roux and Rouanet, 1984; 1988; 2010). For example, in Table 6, the presence of demonstrative pronouns and the presence of BE as a main verb are both assigned the same coordinate on Dimension 1 (0.997). This means that these features co-occur often in the tweets, as can be observed in (1iv) and (2iv). Alternatively, when the distance between the coordinates of the categories of linguistic features is large (i.e. the same coordinate but in opposite quadrants), then this means that these variables are not similarly distributed in the tweets, and hence rarely or never occur with each other in the tweets. For example, in Table 6 the coordinates of the presence of imperatives (-0.997) and the presence of demonstrative pronouns (0.997) on Dimension 1 are far apart – one is negative and the other positive, meaning that these two features never occur with each other in the same tweet in this dataset. This is the largest distance and this indicates that these features distinguish the tweets neatly. Equation 11 demonstrates how the coordinate of a particular category of a linguistic feature is calculated.

When the presence of a linguistic feature occurs in half of the data, the absence of that linguistic feature also occurs in half of the data. When this happens the coordinate of the presence of the linguistic feature is a mirror image of the coordinate of the absence of that linguistic features (one is positive and the other is negative), as this feature distinguishes the tweets. For example, consider the feature BE as a main verb in Table 6. The coordinate of the presence of BE as a main verb (0.997) is a mirror image of the coordinate of the absence of BE as a main verb with an opposite sign (-0.997) and it is the furthest possible distance between two points because BE as a main verb is present across the tweets at the same rate as it is absent (i.e. BE as a main verb is present in 2 tweets and absent in 2 tweets), so this feature neatly distinguishes the four tweets.

However, features are not always equally present and absent across the data. In fact, some features may be present in more than half of the tweets in the dataset, which means that this feature does not distinguish the tweets neatly. As a result, the distance between the presence of a feature and the absence of that feature is not the largest distance (i.e. the same coordinate in opposite quadrants). For example, consider analytic negation in Table 6. Unlike BE as a main verb, the coordinate of the presence of analytic negation (1.107) and the coordinate of the absence of analytic negation (-0.369) are not the same with different signs because analytic negation only

Table 6: The features and their coordinates and contributions for Dimension 1 from the MCA for MDA demonstration

Categories	Dimension 1	
	Coordinate	Contribution
ANALYTIC-NEGATION_A	-0.369	1.271
ANALYTIC-NEGATION_P	1.107	3.812
BE-MAIN-VERB_A	-0.997	6.184
BE-MAIN-VERB_P	0.997	6.184
DEMONSTRATIVE-PRONOUN_A	-0.997	6.184
DEMONSTRATIVE-PRONOUN_P	0.997	6.184
EMOJI_A	-0.296	0.816
EMOJI_P	0.887	2.447
FULLSTOP_A	0.997	6.184
FULLSTOP_P	-0.997	6.184
IMPERATIVE_A	0.997	6.184
IMPERATIVE_P	-0.997	6.184
INITIALMENTION_P	0	0
MULTI-WORD-VERB_A	0.997	6.184
MULTI-WORD-VERB_P	-0.997	6.184
GENERAL-NOUN_A	0.997	6.184
GENERAL-NOUN_P	-0.997	6.184
PREDICATIVE-ADJECTIVE_A	-0.997	6.184
PREDICATIVE-ADJECTIVE_P	0.997	6.184
CONTRACTION_A	-0.369	1.271
CONTRACTION_P	1.107	3.812

occurs in 1 of the four tweets (see Table 4), whereas it is absent in 3 of the 4 tweets. Thus, the absence of analytic negation is more common across the tweets than the presence of analytic negation. This means that the absence of analytic negation will be positioned closer to the features that are shared across the 3 tweets that have the absence of analytic negation, whereas the presence of analytic negation will be positioned further away from the centre, as no other tweets in this dataset share this particular feature. At the same time, the presence of analytic negation will also be positioned closer to the other features that co-occur in this tweet, like the presence of BE as a main verb and demonstrative pronoun (see Table 4).

Table 6 also shows the contributions of each category of a linguistic feature to Dimension 1. The contributions of the categories of linguistic features show which features are most strongly contributing to the dimension. Equation 13 shows how the contributions of categories are calculated. Following Le Roux and Rouanet (2010), the variables contributing above the average contribution should be interpreted, as these represent the most distinguishing patterns of variation. All the contributions for each dimension equal 100 and there are 21 categories of variables in (1iv)-(4iv). Therefore, the average contribution of the categories of features in these examples is 4.76 (100/21=4.76), meaning that anything above this number needs to be interpreted. Because the coordinates indicate the distribution of the categories of features in the tweets, the strongly contributing features with positive coordinates need to be interpreted in opposition to the strongly contributing features with negative coordinates. Table 6 shows that the *presence* of BE as a main verb, demonstrative pronouns, full stops, and predicative adjectives and the *absence* of imperatives, general nouns, and multi-word/phrasal verbs all strongly contribute to the positive side of Dimension 1 as their contributions are above 4.76 and their coordinates are positive. Additionally, these features are also distributed similarly across the tweets, as they all have the same coordinate (0.997). This co-occurrence pattern is in opposition to the *presence* of imperatives, general noun and multi-word/phrasal verbs and the *absence* of BE as a main verb, demonstrative pronouns, full stops, and predicative adjectives, which are all strongly contributing to the negative side of Dimension 1, and also are similarly distributed in the example tweets, as they all have the same negative coordinate (-0.997). Dimension 1 is representing the most important pattern of linguistic co-occurrence in the example tweets. Based on the notion of linguistic co-occurrence (Biber, 1988), these patterns need to be interpreted for their underlying communicative function.

To interpret these patterns, the individual tweets can be examined to view the co-occurring features in context. The MCA assigns the same measures (contributions and coordinates) to each tweet on each Dimension (see Table 7). Similar to the categories of linguistic features, the tweets with high positive and negative coordinates that are most strongly contributing to the Dimension are then interpreted along with the features associated to the corresponding side of the Dimension for their underlying communicative function.

Although this approach factors in the absences of features, which are arguably just as important as what is present, it is hard to interpret the function of the absence of features in the context of tweets as absent features are unobservable. Moreover, features which are absent and are strongly contributing to one side of the dimension tend to also be present and strongly contributing to the other side of the dimension, such as in this demonstration. Thus, to avoid repetition (i.e. discussing the potential function of the absence of features on one side of the dimension and then discussing the function of their presence on the other), the absences of features are ignored and only the observable presences of features are interpreted.

Table 7: The coordinates and contributions of the tweets for Dimension 1 from the demonstration of MCA for MDA

Individuals	Dimension 1	
	Coordinate	**Contribution**
1iv	-0.852	24.848
2iv	-0.852	24.848
3iv	0.758	19.663
4iv	0.946	30.64

Based on this example dataset, positive Dimension 1 is characterised by the presence of demonstrative pronouns, BE as a main verb, and predicative adjectives, which co-occur in examples (3iv) and (4iv). These features are functioning in the Tweets to encode a judgement or evaluation on a previously mentioned event. Alternatively, negative Dimension 1 contains the presence of imperatives, general nouns, full stop and multiword verbs and these are used in (1iv) and (2iv) to instruct the mentioned individual to do something.

To answer research question 1 and research question 2, which aim to describe the major patterns of linguistic variation across Twitter trolling and general English Twitter, this short text version of MDA is applied individually to the corpus of general English tweets and the corpus of trolling tweets in R using the package ‘FactoMineR’ (Husson et al., 2017). The results of these analyses are described in Chapter 5 (for general English Twitter) and Chapter 6 (for Twitter trolling).

4.7.2. Supplementary elements

The set of tweets and linguistic features on which the clouds are constructed are called active individuals and active variables. In addition to active elements, supplementary elements can be included in an MCA to work out the positions of other tweets or the association of other variables to the dimension patterns without affecting the main results of the analysis. Supplementary variables can be qualitative and quantitative. Supplementary qualitative variables are essentially similar to the active variables in that they have categories (Le Roux and Rouanet, 2010). For instance, a supplementary qualitative variable could be the source of the tweet, such as iPhone, Android, Web, iPad, etc., and these can have explanatory power or enrich the interpretation. Clarke and Grieve's (2019) analysis of Donald Trump's tweets, for example, specified the device from which the tweet was sent as a supplementary qualitative variable. They found that the iPhone device was positioned close to the linguistic co-occurrence pattern, which they characterised as a 'campaigning' communicative style. This corresponded to reports suggesting that Donald Trump's campaign team managed his iPhone during the 2016 election (Bulman, 2016), thereby enriching their interpretation. Supplementary categories can be positioned on the cloud of categories using Equation 11.

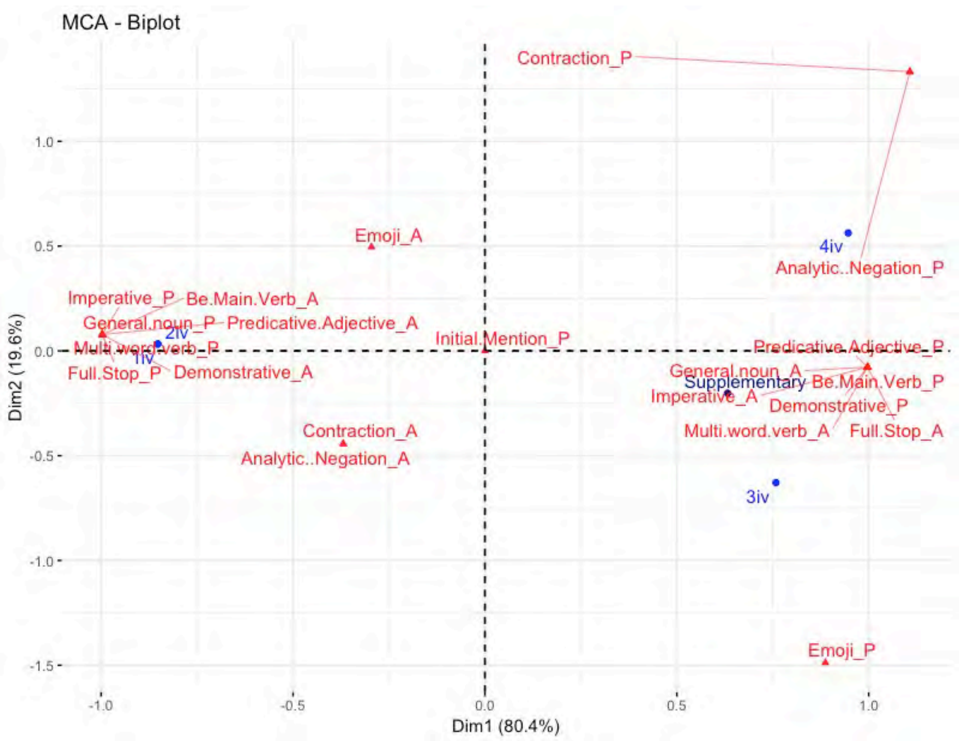
Supplementary quantitative variables could be the age of the Twitter user, the time of day that the tweet was sent, or the length of the tweet in words. Quantitative variables are correlated to the dimension patterns, where a strong correlation might indicate that the quantitative supplementary variable could be used to describe the dimension patterns. In the present dissertation, tweet length in word tokens was included in the trolling/general English tweets by linguistic features data matrices and defined in each individual MCA as a quantitative supplementary variable. The reason for this is because there is the potential problem that text length could confound the analysis because it has not been controlled for in the analysis of the presence/absence of linguistic features.

In particular, the modified version of MDA proposed here does not analyse the relative frequencies of features. The relative frequencies of features are measured in standard MDA as a way to control for texts of different lengths. Measuring the relative frequencies of features, as opposed to their absolute frequency means that texts of different lengths can be compared reliably as the frequencies of features are relative to the length of the text. Thus, by only measuring the presence or absence of features in the short-text version of MDA proposed here, text length is not controlled for and could confound the analysis, as the more words a tweet has the more likely it is to contain a variety of different linguistics features. Defining tweet length as a supplementary variable enables the assessment of the degree to which tweet length is correlated to the results of the analyses.

In addition to supplementary variables, once the clouds of points are constructed, supplementary tweets may also be positioned within the cloud of tweets without contributing to the main results (Le Roux and Rouanet, 2010). This is achieved by measuring the new tweet(s) for the presence or absence of the linguistic features used in the construction of the cloud (i.e. the active variables). Subsequently, the tweet(s) can be located using Equation 10. For example, consider the following supplementary tweet.

(**Supplementary**) This is stupid

Figure 3: Biplot of the cloud of tweets and cloud of categories using MCA-for-MDA on examples (1iv)-(4iv)



This supplementary trolling tweet was analysed for the presence and absence of its linguistic features and combined with the data matrix of examples (1iv)-(4iv) in Table 4. This tweet was defined as a supplementary tweet in the MCA and it was projected onto the cloud of tweets based on its linguistic composition. Figure 3 is a biplot of the cloud of active tweets (1iv)-(4iv) and the cloud of active categories. The supplementary tweet has also been projected onto the biplot. This supplementary tweet is positioned close to the positive side of Dimension 1, representing an evaluative communicative function, and has been positioned closest to example (3iv).

Thus, in a similar way to previous MDA studies that have compared registers and varieties of language to existing dimensions of linguistic variation (e.g. Sardinha, 2014; Jonsson, 2015), the method of supplementary individuals enables new tweets to be geometrically situated with respect to the cloud of tweets without affecting the analysis.

The third research question of this dissertation aims to understand how trolling tweets compare to general English tweets. This dissertation answers this question by comparing the two sets of dimensions and also by comparing trolling tweets with respect to the major patterns of linguistic variation of general English tweets to observe where trolling tweets locate with respect to general English tweets.

To do this, the trolling tweets were measured for their presence or absence of the active linguistic features in the MCA of general English tweets. The trolling tweets were specified as supplementary individuals in the MCA of general English Twitter. The MCA assigned each supplementary trolling tweet a positive or negative coordinate for each dimension of linguistic variation of general English Twitter, which located their position on the cloud of active general English tweets, and indicates how associated they are to particular patterns of co-occurring linguistic features. The overall tendencies of trolling tweets in relation to the dimensions of general English Twitter are described in Chapter 7.

Because each trolling tweet is assigned a coordinate revealing how associated it is to either set of dimensions, the degree of similarity between the dimensions of linguistic variation across the two studies can be assessed. The coordinates of trolling tweets for the dimensions of linguistic variation of general English Twitter were correlated to the coordinates of trolling tweets for the dimensions of linguistic variation of Twitter trolling. A strong correlation would indicate that the dimensions of linguistic variation across the two corpora are the same. The overall result of this comparison is presented at the beginning of Chapter 6 prior to the description of the range of linguistic variation of Twitter trolling.

4.8. Assessing Representativeness

Having illustrated how the range of linguistic variation was identified in the two corpora, it is now possible to demonstrate how the corpora were evaluated for their representativeness. Biber (1993) suggests that the representativeness of a corpus can be evaluated for the extent to which it includes the range of linguistic distributions in the language variety. Biber (1993) proposes that in order to design a representative corpus, a pilot corpus needs to be collected and then this pilot corpus needs to be assessed for the variation in the use of the linguistic structures of interest by conducting a pre-analysis, described below. Based on the internal variation revealed in this pre-analysis, one can subsequently make a calculation about the number of texts needed to achieve representativeness. Essentially, less internal variation requires fewer texts. In one of the first studies to complete this rigorous approach to corpus design, Sardinha (2018) collected a pilot corpus and then tagged the texts using the Biber Tagger and then ran each text through Biber’s Tag Count program, which provided the relative frequencies of all the grammatical and lexical features for each text, and based on these frequencies, it also assigned each text a dimension score for the five major dimensions of linguistic variation in Biber (1988), which revealed how associated each text was to the particular dimensions of linguistic variation of spoken and written English. These scores were used to assess the internal variation by computing the standard deviation of each register on each dimension, as well as the average normalised deviation for each register, and the sum of the average normalised dimensions. These calculations were then used to calculate the size needed for each sub-corpus (cf. Biber, 1993: 255). This approach is useful in cross-register comparisons, such as Sardinha's (2018), which was investigating the range of linguistic variation across various Internet registers, as it is important that a corpus is not biased towards one particular register over the others.

The present dissertation, however, is not conducting a cross-register MDA, and it is not comparing Twitter to Biber's (1988) dimensions of linguistic variation, but rather it is conducting a full MDA of Twitter trolling in order to describe the major patterns of linguistic variation, as well as a full MDA of general English Twitter in order to compare trolling tweets to the major patterns of linguistic variation of general English Twitter. Moreover, the tagger is different in this dissertation, and by extension the linguistic features analysed in this study. Consequently, this dissertation offers a new approach for evaluating representativeness, which is based on the same principle of assessing variation in the range of linguistic distributions, but rather than collect more data or a pilot corpus, this approach instead uses smaller samples of the collected corpus, and instead assesses whether the range of linguistic distributions (of interest to the present dissertation) are stable across the smaller samples.

Specifically, the range of linguistic distributions are measured in this approach by subjecting various smaller random samples of the larger corpus to short-text MDA (described in section 4.6). Each smaller random sample of tweets is measured for the presence or absence of the linguistic features that occur in more than 5 percent of all of the tweets in the full corpus. The presence and absence of these linguistic features in the smaller random samples are recorded in separate categorical data matrices. These data matrices recording the presence or absence of the linguistic features occurring in the smaller random samples are then separately subjected to Multiple Correspondence Analysis (MCA), which produces a series of dimensions representing the most common patterns of variation across the particular smaller random sample of tweets. For each dimension in each random sample, the MCA assigns each category (presence or absence) of a linguistic feature a positive or negative coordinate and a value indicating its contribution to the dimension. The coordinates and contributions of the linguistic features in each dimension reveal which features co-occur together most often across the sample of tweets. Thus, the coordinates and contributions of linguistic features in each dimension in each sample represent the distribution of the linguistic features across the tweets in the sample and reflect the most common patterns of linguistic variation in that particular sample.

To assess the variation in the linguistic distributions across these smaller random samples of the data, the coordinates and contributions of all the linguistic features in each dimension in one sample are correlated to the other. The coordinates of the linguistic features and the contributions of the linguistic features in each dimension in each sample can be correlated to each other because each sample is analysed for the same feature set. In other words, there is no disparity between the actual variables in each MCA of the samples, the only possible disparity between the variables in each sample are the coordinates and contributions assigned to them by the MCA, and these are dependent on their distribution in the tweets in the smaller random samples. Thus, correlating the coordinates and contribution of the linguistic features across the dimensions of one sample with the coordinates and contributions of the linguistic features across the dimensions of the other sample enables the assessment of variation in the linguistic distributions across the random samples. Strong correlations between the dimensions in each sample are indicative of a lack of variation between the two samples in terms of the distribution of linguistic features in the tweets, whilst weak correlations suggest greater variation across the samples in terms of the distribution of the linguistic features. A lack of variation indicates that the range of linguistic distributions across these samples is represented in these samples. The range of linguistic distributions are therefore the dimensions of linguistic variation revealed in the short text MDA of the samples, and the representativeness of the range of linguistic distributions in the corpus is therefore the stability of the dimensions across these smaller samples. Dimensions of linguistic variation are stable when they are correlated in smaller random samples of the larger corpus, and thus display little variation, despite having fewer and random chunks of the overall data. Essentially, if the full range of linguistic distributions is found in smaller random samples of the larger corpus and if these linguistic distributions are relatively stable in other smaller random samples, then it can therefore be argued that the original corpus is representative of the range of linguistic distributions in that variety. This is because if more data does not lead to any more important, new or different linguistic distributions (i.e. patterns of linguistic variation), then the data has arguably reached a point of saturation. Additionally, if the dimensions of linguistic variation are stable when compared with other smaller samples of tweets, then it can be argued that these samples are representative of the range of variation, suggesting that no more data is needed for interpreting the major dimensions of linguistic variation.

In addition to the linguistic features, the MCA also assigns each tweet in the sample a positive or negative coordinate and a value indicating its contribution to the dimension. The coordinates and contributions of the tweets cannot be correlated across samples as they are not the same, so there is no basis for correlation. Moreover, the tweets in each sample are purposely not the same in order to assess the range of linguistic distributions in different random smaller samples of the corpus.

To implement this on the general Twitter corpus, smaller random samples of tweets were extracted from the general Twitter corpus. Table 8 presents the different samples and their sizes. For sample size (n), two random samples of that size were extracted without replacement. For example, the first sample size is 500 tweets, and so two lots of 500 random tweet samples were extracted, and these 500 tweet samples are completely distinct in that the tweets found in the first 500-tweet sample are not in the second 500-tweet sample. This was repeated 5 times for each sample size. Thus, there were 10 samples for each sample size. These smaller samples of tweets from the larger corpus were tagged for the presence or absence of the 63 linguistic features, and this information was recorded in separate data matrices for each sample of tweets (sample size (*n*) tweets by 63 features).

Each matrix was subsequently subjected to MCA and then the coordinates and contributions of the categories of the linguistic features for each of the major dimensions were extracted for each sample. As mentioned, coordinates reflect the nature of the association between features in terms of proximity, where features that are more associated or are distributed in similar ways in the data will have coordinates closer to each other (Le Roux and Rouanet, 1984; 1988; 2010). Contributions reflect which features are the most important contributors to the particular dimension. Using these measures, dimension stability was assessed by correlating the contributions of the linguistic features in one sample with the contributions of the linguistic features in the other same-sized sample across the five most important dimensions. Additionally, the coordinates of the linguistic features for each of these dimensions in one sample were correlated with the coordinates of the linguistic features for each dimension in the other same-sized sample. This process was repeated for the 5 sets of same-sized distinct samples.

The results of these correlations can be found in Appendix 2 and a summary of the strongest pair correlations for each set of samples is also provided in Table 8. Table 8 shows that the first four dimensions of linguistic variation are relatively stable in different samples of 2,000 tweets and beyond and that the first five dimensions are strongly correlated in different samples of 6,000 tweets on both coordinates and contributions, although occasionally Dimension 3 in one sample is correlated to the other sample’s Dimension 4 and vice versa (see Table 9), suggesting that the particular patterns of linguistic variation in either dimension is more common in one sample than the other. Table 9, for example, shows this pattern more clearly.

Table 8: Summary of the strongest dimension pair correlations for coordinates and contributions for each set of samples of General English tweets

Sample	Sample size	Correlated (*r*) Dimension (D*n)* coordinates	Correlated (*r*) Dimension (*Dn*) contributions
S1 and S2	500	D1S1-D1S2 = 0.96 D2S1- D2S2 = -0.86 D3S1-D3S2 = 0.66	D1S1-D1S2 = 0.88 D2S1- D2S2 = 0.92
S3 and S4		D1S3-D1S4 = 0.95 D2S3- D2S4 = -0.92	D1S3-D1S4 = 0.91 D2S3- D2S4 = 0.92
S5 and S6		D1S5-D1S6 = 0.97 D2S5- D2S6 = -0.91	D1S5-D1S6 = 0.94 D2S5- D2S6 = 0.86 D4S5-D3S6 = 0.69
S7 and S8		D1S7-D1S8 = 0.96 D2S7- D2S8 = -0.85	D1S7-D1S8 = 0.9 D2S7- D2S8 = 0.92
S9 and S10		D1S9-D1S10 = 0.97 D2S9- D2S10 = 0.88	D1S5-D1S6 = 0.9 D2S5- D2S6 = 0.9 D4S5-D3S6 = 0.7
S11 and S12	1000	D1S11-D1S11 = 0.97 D2S12- D2S12 = -0.95 D3S11-D3S12 = 0.67 D4S11-D4S12 = 0.57	D1S11-D1S11 = 0.95 D2S12- D2S12 = 0.95 D3S11-D3S12 = 0.76 D4S11-D4S12 = 0.6
S13 and S14		D1S13-D1S14 = 0.97 D2S13- D2S14 = 0.94 D4S13-D4S14 = 0.69	D1S13-D1S14 = 0.93 D2S13- D2S14 = 0.96 D3S13-D3S14 = 0.65 D4S13-D4S14 = 0.6
S15 and S16		D1S15-D1S16 = 0.99 D2S15- D2S16 = 0.91 D3S15-D4S16 = 0.67	D1S15-D1S16 = 0.97 D2S15- D2S16 = 0.91
S17 and S18		D1S17-D1S18 = 0.96 D2S17- D2S18 = 0.95	D1S17-D1S18 = 0.95 D2S17- D2S18 = 0.95 D4S17-D4S18 = 0.64
S19 and S20		D1S19-D1S20 = 0.99 D2S19- D2S20 = 0.97	D1S19-D1S20 = 0.97 D2S19- D2S20 = 0.97
S21 and S22	2000	D1S21-D1S22 = 0.99 D2S21- D2S22 = 0.97 D3S21-D4S22 = 0.79 D4S21-D3S22 = 0.67	D1S21-D1S22 = 0.97 D2S21- D2S22 = 0.95 D3S21-D4S22 = 0.67 D4S21-D3S22 = 0.83
S23 and S24		D1S23-D1S24 = 0.99 D2S23- D2S24 = 0.98 D3S23-D5S24 = 0.55 D4S23-D3S24 = 0.6	D1S23-D1S24 = 0.97 D2S23- D2S24 = 0.98 D3S23-D4S24 = 0.66 D4S23-D3S24 = 0.7
S25 and S26		D1S25-D1S26 = 0.99 D2S25- D2S26 = 0.97	D1S25-D1S26 = 0.97 D2S25- D2S26 = 0.96

		D3S25-D3S26 = 0.55	D3S25-D3S26 = 0.55
		D4S25-D4S26 = 0.77	D4S25-D3S26 = 0.65
		D5S25-D5S26 = 0.78	D5S25-D5S26 = 0.73
S27 and S28		D1S27-D1S28 = 0.99	D1S27-D1S28 = 0.98
		D2S27- D2S28 = 0.97	D2S27- D2S28 = 0.97
		D3S27-D3S28 = 0.82	D3S27-D3S28 = 0.88
		D4S27-D4S28 = 0.76	D4S27-D4S28 = 0.71
S29 and S30		D1S29-D1S30 = 0.99	D1S29-D1S30 = 0.98
		D2S29- D2S30 = 0.97	D2S29- D2S30 = 0.98
		D3S29-D3S30 = 0.71	D3S29-D3S30 = 0.78
		D4S29-D4S30 = 0.6	D4S29-D4S30 = 0.62
S31 and S32	3000	D1S31-D1S32 = 0.99	D1S31-D1S32 = 0.98
		D2S31- D2S32 = 0.98	D2S31- D2S32 = 0.98
		D3S31-D4S32 = 0.89	D3S31-D4S32 = 0.94
		D4S31-D3S32 = 0.87	D4S31-D3S32 = 0.9
		D5S31-D5S32 = 0.72	D5S31-D5S32 = 0.68
S33 and S34		D1S33-D1S34 = 0.99	D1S33-D1S34 = 0.98
		D2S33- D2S34 = 0.98	D2S33- D2S34 = 0.98
		D3S33-D4S34 = 0.72	D3S33-D3S34 = 0.77
		D4S33-D3S34 = 0.67	
S35 and S36		D1S35-D1S36 = 0.99	D1S35-D1S36 = 0.98
		D2S35- D2S36 = 0.99	D2S35- D2S36 = 0.98
		D3S35-D3S36 = -0.75	D3S35-D3S36 = 0.81
		D4S35-D4S36 = 0.68	D4S35-D4S36 = 0.72
S37 and S38		D1S37-D1S38 = 0.99	D1S37-D1S38 = 0.98
		D2S37- D2S38 = 0.98	D2S37- D2S38 = 0.98
		D3S37-D4S38 = 0.88	D3S37-D4S38 = 0.9
		D4S37-D3S38 = 0.85	D4S37-D3S38 = 0.92
S39 and S40		D1S39-D1S40 = 0.99	D1S39-D1S40 = 0.99
		D2S39- D2S40 = 0.97	D2S39- D2S40 = 0.97
		D3S39-D3S40 = 0.92	D3S39-D3S40 = 0.96
		D4S39-D4S40 = 0.8	D4S39-D4S40 = 0.87
			D5S39-D5S40 = 0.78
S41 and S42	4000	D1S41-D1S42 = 0.99	D1S41-D1S42 = 0.99
		D2S41- D2S42 = 0.99	D2S41- D2S42 = 0.99
		D3S41-D4S42 = 0.91	D3S41-D4S42 = 0.89
		D4S41-D3S42 = 0.91	D4S41-D3S42 = 0.94
		D5S41-D5S42 = 0.71	D5S41-D5S42 = 0.76
S43 and S44		D1S43-D1S44 = 0.99	D1S43-D1S44 = 0.99
		D2S43- D2S44 = 0.99	D2S43- D2S44 = 0.99
		D3S43-D4S44 = 0.87	D3S43-D4S44 = 0.93
		D4S43-D3S44 = -0.9	D4S43-D3S44 = 0.92
		D5S43-D5S44 = 0.78	D5S43-D5S44 = 0.79
S45 and S46		D1S45-D1S46 = 0.99	D1S45-D1S46 = 0.99
		D2S45- D2S46 = 0.98	D2S45- D2S46 = 0.99
		D3S45-D4S46 = 0.68	D3S45-D3S46 = 0.76
		D4S45-D3S46 = -0.73	D4S45-D4S46 = 0.7
		D5S45-D5S46 = 0.69	D5S45-D5S46 = 0.86
S47 and S48		D1S47-D1S48 = 0.99	D1S47-D1S48 = 0.99
		D2S47- D2S48 = 0.99	D2S47- D2S48 = 0.98
		D3S47-D3S48 = 0.8	D3S47-D3S48 = 0.8
		D4S47-D4S48 = 0.89	D4S47-D4S48 = 0.83
S49 and S50		D1S49-D1S50 = 0.99	D1S49-D1S50 = 0.99
		D2S49- D2S50 = 0.99	D2S49- D2S50 = 0.99
		D3S49-D4S50 = 0.75	D3S49-D4S50 = 0.75
		D4S49-D3S50 = 0.74	D4S49-D3S50 = 0.62
		D5S49-D5S50 = 0.69	
S51 and S52	5000	D1S51-D1S52 = 0.99	D1S51-D1S52 = 0.99

		D2S51- D2S52 = 0.99	D2S51- D2S52 = 0.99
		D3S51-D4S52 = 0.91	D3S51-D4S52 = 0.96
		D4S51-D3S52 = 0.87	D4S51-D3S52 = 0.91
		D5S51-D5S52 = 0.93	D5S51-D5S52 = 0.94
S53 and S54		D1S53-D1S54 = 0.99	D1S53-D1S54 = 0.99
		D2S53- D2S54 = 0.99	D2S53- D2S54 = 0.99
		D3S53-D4S54 = 0.77	D3S53-D4S54 = 0.67
		D4S53-D3S54 = 0.79	D4S53-D3S54 = 0.78
		D5S53-D5S54 = 0.6	D5S53-D5S54 = 0.73
S55 and S56		D1S55-D1S56 = 0.99	D1S55-D1S56 = 0.99
		D2S55- D2S56 = 0.99	D2S55- D2S56 = 0.99
		D3S55-D3S56 = 0.9	D3S55-D3S56 = 0.96
		D4S55-D4S56 = 0.94	D4S55-D4S56 = 0.93
		D5S55-D5S56 = 0.79	D5S55-D5S56 = 0.74
S57 and S58		D1S57-D1S58 = 0.99	D1S57-D1S58 = 0.99
		D2S57- D2S58 = 0.98	D2S57- D2S58 = 0.98
		D3S57-D4S58 = -0.71	D3S57-D3S58 = 0.68
		D4S57-D3S58 = 0.85	D5S57-D5S58 = 0.96
		D5S57-D5S58 = 0.92	
S59 and S60		D1S59-D1S60 = 0.99	D1S59-D1S60 = 0.99
		D2S59- D2S60 = 0.99	D2S59- D2S60 = 0.99
		D3S59-D4S60 = 0.93	D3S59-D4S60 = 0.93
		D4S59-D3S60 = 0.87	D4S59-D3S60 = 0.9
S61 and S62	6000	D1S61-D1S62 = 0.99	D1S61-D1S62 = 0.99
		D2S61- D2S62 = 0.99	D2S61- D2S62 = 0.99
		D3S61-D3S62 = 0.89	D3S61-D3S62 = 0.92
		D4S61-D4S62 = 0.9	D4S61-D4S62 = 0.91
S63 and S64		D1S63-D1S64 = 0.99	D1S63-D1S64 = 0.99
		D2S63- D2S64 = 0.99	D2S63- D2S64 = 0.98
		D3S63-D4S64 = 0.91	D3S63-D4S64 = 0.9
		D4S63-D3S64 = 0.92	D4S63-D3S64 = 0.95
		D5S63-D5S64 = 0.86	D5S63-D5S64 = 0.92
S65 and S66		D1S65-D1S66 = 0.99	D1S65-D1S66 = 0.99
		D2S65- D2S66 = 0.99	D2S65- D2S66 = 0.99
		D3S65-D3S66 = 0.77	D3S65-D3S66 = 0.61
		D4S65-D4S66 = -0.8	D4S65-D4S66 = 0.74
		D5S65-D5S66 = 0.91	D5S65-D5S66 = 0.92
S67 and S68		D1S67-D1S68 = 0.99	D1S67-D1S68 = 0.99
		D2S67- D2S68 = 0.99	D2S67- D2S68 = 0.99
		D3S67-D3S68 = 0.92	D3S67-D3S68 = 0.95
		D4S67-D4S68 = -0.96	D4S67-D4S68 = 0.97
		D5S67-D5S68 = 0.75	D5S67-D5S68 = 0.79
S69 and S70		D1S69-D1S70 = 0.99	D1S69-D1S70 = 0.99
		D2S69- D2S70 = 0.99	D2S69- D2S70 = 0.99
		D3S69-D4S670 = 0.86	D3S69-D4S670 = 0.84
		D4S69-D3S70 = 0.88	D4S69-D3S70 = 0.87
		D5S69-D5S70 = 0.94	D5S69-D5S70 = 0.97

Table 9 is the combined coordinates and contribution correlation matrices of sample 69 and 70, which are 6000-word samples. Importantly, Tables 8 and 9 indicate that by 6,000 tweets there is little variation amongst the distribution of linguistic features in the five most important dimensions.

Specifically, the dimensions of linguistic features in 6,000-tweet samples are exceptionally stable in terms of their association/proximity to each other (i.e. how often they co-occur in the tweets), as evidenced in the coordinates, and in terms of their importance to the overall dimension pattern, as evidenced in the contributions.

Table 9: Correlation matrices of the 6000-word sample's (69 and 70) coordinates and contributions

Correlation of Coordinates 6000-word sample 69 and 70					
S69/S70	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.99	-0.12	-0.08	0.07	0.04
Dim_2	-0.12	0.99	0.11	-0.05	0.02
Dim_3	0.04	0.03	0.3	0.88	0.08
Dim_4	-0.1	0.06	0.86	-0.48	0.01
Dim_5	0.03	-0.003	0.1	0.04	0.94
Correlation of Contribution 6000-word sample 69 and 70					
S69/S70	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.99	-0.06	0.18	0.37	-0.05
Dim_2	-0.06	0.99	0.23	-0.02	0.08
Dim_3	0.37	-0.002	0.35	0.87	-0.01
Dim_4	0.22	0.33	0.84	0.39	0.17
Dim_5	-0.05	0.05	0.16	-0.01	0.97

The same process was completed on the Twitter trolling corpus, although for this analysis the feature set occurring in more than 5% of these tweets was used (see Appendix 1 and 2). Separate data matrices of sample size (n) tweets by 69 linguistic features were subjected to MCA and the results (the coordinates and contributions of linguistic features) were correlated. Table 10 presents a summary of the results of comparing equal-sized random and unique samples of the Twitter trolling corpus. Table 10 shows that at 2000 tweets at least the first four dimensions are relatively stable, which indicates that these linguistic distributions are relatively stable. Dimension 5, however, is a little bit more precarious in these smaller samples.

Table 10: Summary of the strongest dimension pair correlations for coordinates and contributions for each set of samples for trolling tweets

Troll Sample	Sample size	Correlated (*r*) Dimension (D*n*) coordinates	Correlated (*r*) Dimension (*Dn*) contributions
S1 and S2	500	D1S1-D1S2 = 0.96 D2S1- D2S2 = -0.72	D1S1-D1S2 = 0.96 D2S1- D2S2 = 0.72 D3S1-D3S2 = 0.68
S3 and S4		D1S3-D1S4 = 0.97 D2S3- D2S4 = -0.73 D3S3-D4S4 = 0.59 D4S3-D3S4 = -0.63	D1S3-D1S4 = 0.92 D2S3- D2S4 = 0.59 D3S3-D4S4 = 0.66 D4S3-D3S4 = 0.62
S5 and S6		D1S5-D1S6 = 0.97 D2S5- D2S6 = -0.65 D3S5-D3S6 = 0.54	D1S5-D1S6 = 0.92 D2S5- D2S6 = 0.61 D3S5-D3S6 = 0.72
S7 and S8		D1S7-D1S8 = 0.98 D2S7- D2S8 = -0.66	D1S7-D1S8 = 0.92 D2S7- D2S8 = 0.62
S9 and S10		D1S9-D1S10 = 0.98 D2S9- D2S10 = 0.77	D1S9-D1S10 = 0.94 D2S9- D2S10 = 0.65
S11 and S12	1000	D1S11-D1S12 = 0.98 D2S11- D2S12 = 0.86 D3S11-D3S12 = 0.68	D1S11-D1S12 = 0.93 D2S11- D2S12 = 0.82 D3S11-D3S12 = 0.78
S13 and S14		D1S13-D1S14 = 0.99 D2S13- D2S14 = 0.86 D3S13-D3S14 = 0.76	D1S13-D1S14 = 0.97 D2S13- D2S14 = 0.75 D3S13-D3S14 = 0.86
S15 and S16		D1S15-D1S16 = 0.99 D2S15- D2S16 = 0.85 D3S15-D3S16 = 0.53	D1S15-D1S16 = 0.97 D2S15- D2S16 = 0.79 D3S15-D3S16 = 0.88
S17 and S18		D1S17-D1S18 = 0.99 D2S17- D2S18 = 0.86 D3S17-D3S18 = 0.52 D4S17-D4S18 = 0.59	D1S17-D1S18 = 0.95 D2S17- D2S18 = 0.76 D3S17-D3S18 = 0.71
S19 and S20		D1S19-D1S20 = 0.99 D2S19- D2S20 = 0.87 D3S19-D3S20 = 0.69 D4S19-D4S20 = 0.52	D1S19-D1S20 = 0.97 D2S19- D2S20 = 0.78 D3S19-D3S20 = 0.85
S21 and S22	2000	D1S21-D1S22 = 0.99 D2S21- D2S22 = 0.89 D3S21-D3S22 = 0.79 D4S21-D4S22 = 0.55	D1S21-D1S22 = 0.98 D2S21- D2S22 = 0.85 D3S21-D3S22 = 0.87 D4S21-D4S22 = 0.5
S23 and S24		D1S23-D1S24 = 0.99 D2S23- D2S24 = 0.92 D3S23-D3S24 = 0.8 D4S23-D4S24 = 0.78	D1S23-D1S24 = 0.98 D2S23- D2S24 = 0.78 D3S23-D3S24 = 0.94 D4S23-D4S24 = 0.65
S25 and S26		D1S25-D1S26 = 0.99 D2S25- D2S26 = 0.93 D3S25-D3S26 = 0.79	D1S25-D1S26 = 0.98 D2S25- D2S26 = 0.89 D3S25-D3S26 = 0.9 D5S25-D4S26 = 0.62
S27 and S28		D1S27-D1S28 = 0.99	D1S27-D1S28 = 0.98

	D2S27- D2S28 = 0.92	D2S27- D2S28 = 0.9
	D3S27-D3S28 = 0.78	D3S27-D3S28 = 0.93
	D4S27-D4S28 = 0.67	D4S27-D4S28 = 0.7
		D527-D528 = 0.67
S29 and S30	D1S29-D1S30 = 0.99	D1S29-D1S30 = 0.98
	D2S29- D2S30 = 0.92	D2S29- D2S30 = 0.93
	D3S29-D3S30 = 0.79	D3S29-D3S30 = 0.92
	D4S29-D4S30 = 0.71	D4S29-D4S30 =0.71
		D5S29-D530 = 0.54

There are at least two possible steps that can be taken at this stage. The first is that it can be argued that the corpus is representative of the linguistic distributions found in the first four major dimensions, and thus the analysis should go no further than interpreting these dimensions. Alternatively, if one desired to interpret Dimension 5 and conclude that the corpus was representative of the first five major dimensions of linguistic variation, then this should be treated as a pilot corpus, and like Biber’s (1993) suggestion, the next step would be to collect more cases of trolling which display these linguistic co-occurrence patterns, and then reassess the internal variation. Rather than collect more data, this thesis instead interprets Dimension 1 to 4 as being representative of the major dimensions of linguistic variation in Twitter trolling. Interpretations of further dimensions of linguistic variation should therefore be viewed as potentially non-representative of the wider trolling repertoire on Twitter, although importantly, given that the correlations were made on smaller samples of the larger corpus, it might be possible that the larger corpus might present further stability on further dimensions, although the veracity of this statement would require more trolling data.

4.9. Summary

The present chapter has described the steps taken to answer the research questions. This chapter has introduced a bespoke grammatical tagger for tagging tweets according to the MDA feature set and Twitter and CMC-specific features. Moreover, this chapter introduces a new modified version of Multi-Dimensional Analysis for the analysis of short texts. Finally, this chapter introduced a new way for assessing representativeness in terms of dimension stability. With access to the data, grammatical tagger and the feature counter, it is possible to replicate this following these steps and this will return the same results. Chapters 5, 6 and 7 present the results of the analyses.

5. Dimensions of General Twitter

To find the major dimensions of linguistic variation of general English Twitter, I used the short text version of MDA described in section 4.7 on the 13,879 general English tweets measured for the presence or absence of 63 linguistic features that occurred in more than 5% of the tweets (see Appendix 1). The MCA returned 63 dimensions ($L \leq 126\ categories - 63\ linguistic\ fea\square ures = 63$). For each dimension, each category (presence or absence) of all 63 linguistic features was assigned a positive or negative coordinate and a value indicating their contribution to the dimension. Additionally, for each dimension each tweet was assigned a positive or negative coordinate and a value indicating its contribution to the dimension. I extracted the first five dimensions, as these were readily interpretable, and based on the modified rates of the eigenvalues (see section 4.7.1), these first five dimensions explain a large proportion of the variance (see Table 11). For each dimension, I interpreted the features that contributed above the average contribution $\left(\frac{100}{63\ linguistic\ features\ x\ 2\ categories} = \frac{100}{126} = \ 0.79\right)$ and the strongly contributing tweets. Those features and tweets contributing the most with positive coordinates were interpreted in opposition to the features and tweets with negative coordinates for the underlying communicative function. Each dimension is presented below in a separate section. Whilst the absences of features are displayed in each dimension's feature table, only the presences of features are interpreted (see section 4.7.1). More examples for each dimension can be found in Appendix 3.

Table 11: Variances of Dimensions (eigenvalues and modified rates)

Dimension	1	2	3	4	5
Eigenvalue	0.108	0.048	0.027	0.027	0.023
Modified Rates	0.85	0.10	0.01	0.01	0.005

5.1. Dimension 1: Length

The linguistic features most strongly contributing to Dimension 1 are presented in Table 12. This table shows that positive Dimension 1 is characterised by the *presence* of 40 linguistic features, whereas negative Dimension 1 has the absence of 11 features. These 11 features are some of the most frequently occurring features in the whole corpus, which suggests that their absence is highly unusual. For instance, *Other_Nouns* refers to general nouns, and *Other_Verbs* refers to general verbs that are not tagged as any of the specialised noun or verb classes. These features occur in 77 percent and 58 percent of tweets in this corpus, respectively.

The opposition between the *presence* of linguistic features with the *absence* of linguistic features is not only observed in these strongly contributing features, but it is also reflected further down the dimension, where *all* linguistic features are present on the positive side, except for *URLs* and *emojis*. This suggests that Dimension 1 is reflecting variation in tweet length, as typically the more words a tweet has the more likely it is to have the presence of numerous linguistic features, as opposed to shorter tweets, which will more likely have the absence of features.

Table 12: The linguistic features strongly contributing to Dimension 1 (coordinate, contribution)

1	**+**	**Present features**: Attributive_Adjective (0.388; 1.042), Other_Verb (0.431; 1.621), Second_Person_Pronoun (0.504; 0.864), Preposition (0.504; 1.91), First_Person_Pronoun (0.539; 1.667), Other_Adverb (0.675; 2.185), Possession (0.701; 1.707), Past_Tense_Verb (0.716; 1.905), Full_Stop (0.723; 2.473), Amplifier (0.731; 0.819), Third_Person_Singular_Verb (0.733; 2.195), Subject_Pronoun (0.739; 2.948), Nominalisation (0.752; 1.005), Definite_Article (0.753; 2.311), Indefinite_Article (0.779; 1.934), Stative_Form (0.786; 2.799), Phrasal_Verb (0.798; 0.943), WH-Word (0.822; 1.057), Object_Pronoun (0.825; 1.149), Comma (0.858; 2.228), Public_Verb (0.859; 1.357), Demonstrative_Pronoun (0.86; 0.965), Coordinating_Conjunct (0.918; 2.725), Contracted_Forms (0.932; 2.003), Modal_Possibility (0.944; 1.055), Predicative_Adjective (0.957; 1.94), Infinitive (0.998; 2.206), IT (1.034; 2.071), Progressive (1.041; 1.175), Private_Verb (1.056; 2.13), Third_Person_Pronoun (1.057; 2.547), Analytic_Negation (1.059; 2.541), Modal_Prediction (1.092; 1.439), Indefinite_Pronoun (1.099; 1.268), HAVE_Main_Verb (1.131; 0.979), General_Subordinator (1.198; 2.947), Passive (1.203; 1.09), Auxiliary_DO (1.24; 1.536), Contrastive_Conjunct (1.251; 1.609), Complementation (1.337; 1.597)
	-	**Absent features**: Other_Noun (-0.702; 1.659), Other_Verb (-0.632; 2.38), Preposition (-0.528; 2.001), Subject_Pronoun (-0.43; 1.716), Stative_Form (-0.35; 1.248), Attributive_Adjective (-0.345; 0.927), First_Person_Pronoun (-0.345; 1.068), Full_Stop (-0.343; 1.174), Other_Adverb (-0.327; 1.057), Definite_Article (-0.289; 0.888), Third_Person_Singular_Verb (-0.282; 0.845).

Additionally, the opposition between the presence of all features except for *URLs* and *emojis* also suggests that these features can often stand alone in tweets or at least often occur in the shortest tweets. These interpretations are supported by the tweets most strongly associated with Dimension 1, presented in Table 13.

Tweets associated with the positive side (Examples 1-5) are considerably longer than those associated with the negative side (Examples

6-10). Example 1, for instance, contains 47 word tokens, whereas Example 6 contains 1 word token, an *emoji* and a *URL*.

Table 13: Examples of the top 5 tweets most associated with positive and negative Dimension 1 according to the strongest contributions and the highest coordinates

	Tweet	Coord	Contrib
1	@realDonaldTrump How can Obama have photos of something you implemented a few days ago? I'm sure he has seen them now if that's what you mean. He's not really in a position to do anything about it, but sure, shuffle blame you insecure coward	1.289	0.111
2	You know the more i think about X's death the more it gets me fucked up. This dude was getting better. He was making an effort to do better. He didn't get a chance. He was MY AGE, probably seen more fucked up shit than any of us, and he was killed for a fucking vutton bag.	1.195	0.095
3	Ray pernah bilang, "You are too mature for Wattpad!" I'd like to believe that it's true, but then again, the biggest question remains: if I couldn't be noticed in digital realm such Wattpad, how would I expect to be considered in the real world?	1.189	0.094
4	@wittymittie for how much I actually tried to read up on this before voting, THAT part I didn't hear.It's absurd, of course, and wouldn't have gotten my vote lmao but I'm just shocked it's a real thing.	1.186	0.094
5	@McoolFionn @zenabby1 Now you're not thinking straight. You see, one have to keep going, in spite of any and all perils, until you reach a country that has sufficient welfare. This "safety" concept of your is irrelevant.At least, this is what I believe goes through their heads, if anything at all	1.181	0.093
6	This 👇 https://t.co/xPBHPbLE2v	-0.481	0.015
7	□ goodnight□goodnighttttt□ https://t.co/MGjWbc4ZRq	-0.479	0.015
8	Congrats 👍👏https://https://t.co/djNZb4EIAW	-0.479	0.015
9	Hello 👋🙌https://t.co/0f0DJ0KWWE	-0.479	0.015
10	Jungkook-ahh😍💞💞💞 https://t.co/Mv9IXsjwqG	-0.479	0.015

Examples 6-10 mainly consist of *URLs* and *emojis*, indicating that they can occur independently. The reason for this may be because emojis and URLs are linguistic units that do not necessarily require other linguistic units to make meaning. For example, the *URLs* in these tweets are links to photos, other Twitter statuses, websites, news articles, and gifs, and these can be shared without other linguistic units and still make meaning. *Emojis* often encode a reaction, and like body language or facial expression, are able to similarly make meaning independently from other linguistic features.

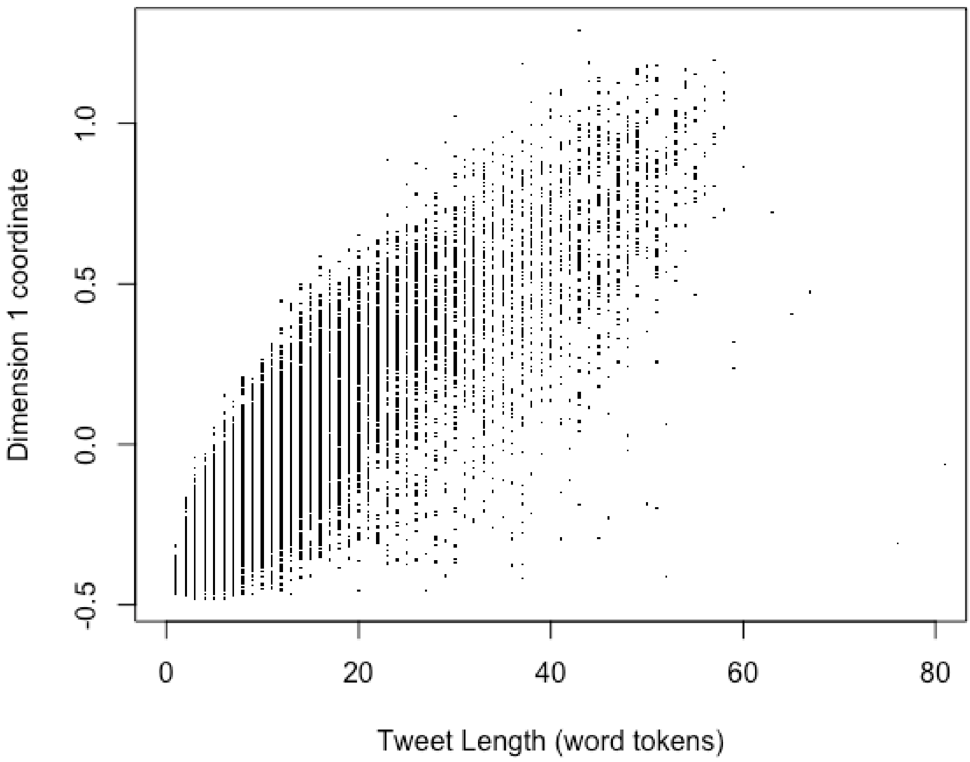
As mentioned in section 4.7.2, I included tweet length as a supplementary quantitative variable as a way to assess whether tweet length has confounded the analysis. Supplementary variables can be used in order to assess the degree to which the variation across the dimensions is correlated to the variable without affecting the main results. Table 14 presents the correlation between tweet coordinates and tweet length for each dimension, showing that Dimension 1 is strongly positively correlated to tweet length, thereby supporting the interpretation that Dimension 1 is explaining text length. The slight positive correlation of tweet length to Dimension 2 will be discussed in the next section.

Table 14: The correlation between tweet coordinates and tweet length for each dimension

	Dim.1	**Dim.2**	**Dim.3**	**Dim.4**	**Dim.5**	**Dim.6**
Tweet Length (word tokens)	0.843	0.281	0.004	-0.015	0.039	-0.041

Figure 4 also displays this correlation, showing that the tweets' Dimension 1 coordinates tend to rise with text length, although the rise begins to slow after 20 words. This flattening occurs at least in part because the likelihood of new grammatical forms occurring for the first time decreases as the length of tweet increases.

Figure 4 General tweets' Dimension 1 coordinate correlated to tweet length.



The strong correlation of tweet length to Dimension 1 is largely because the length of the tweet is the strongest influence on the presence or absence of features. In other words, as tweets get longer, they tend to have the presence of more linguistic features because without words, it is not possible to have the presence of features.

The strong correlation to tweet length is also because tweet length is not controlled for in the short text version of MDA, which measures the presence or absence of linguistic features, as opposed to their frequency relative to the length of the text, like in standard MDA. Relative frequencies of features are analysed in standard MDA to compare texts of different lengths reliably. Given that only analysing for the presence or absence of features does not control for text length, it is not surprising tweet length influences the results. However, apart from a slight correlation to Dimension 2, tweet length is only strongly correlated to Dimension 1, suggesting that tweet length has been largely controlled for in the first dimension. Because of this strong correlation, Dimension 1 excluded from further linguistic interpretation.

5.2. Dimension 2: Informational versus Interactive

The linguistic features that are strongly contributing to Dimension 2 are presented in Table 15. Table 15 shows that positive Dimension 2 is characterised by various noun types, such as *general nouns*, *nominalisations, proper nouns,* and *numeral nouns,* suggesting that the tweets associated with this side of the dimension have numerous referents. *Non-initial mentioning* is another noun type, enabling another Twitter user to be referenced in the third person, as opposed to interacting with them directly (Honeycutt and Herring, 2009). Additionally, there are several noun modifiers associated with positive Dimension 2, enabling the concise and efficient integration of information (Biber, 1988), such as *definite articles, numeral determiners* and *attributive adjectives*, which are used for pre-nominal modification. Prepositions are also associated with positive Dimension 2, and are often used in complex noun phrases as post-nominal modification. A high degree of noun types and noun modifiers are associated with texts that have a high informational focus (Biber, 1988).

Table 15: The linguistic features strongly contributing to positive and negative Dimension 2 (coordinate; contribution)

2	+	**Present features**: Other_Noun (0.199; 1.008), Attributive_Adjective (0.277; 1.193), Definite_Article (0.341; 1.067), Preposition (0.389; 2.563), Imperative (0.496; 0.99), Proper_Noun (0.524; 4.147), URL (0.637; 5.813), Capitalisation (0.663; 3.469), Nominalisation (0.703; 1.977), Numeral_Noun (0.856; 2.422), Numeral_Determiner (0.988; 3.001), Non-Initial_Mention (1.043; 3.695), Hashtag (1.117; 6.62), Colon (1.5; 4.927).
		Absent features: Subject_Pronoun (0.249; 1.292), Initial_Mention (0.387; 3.175),
	-	**Present features**: Initial_Mention (-0.688; 5.638), Auxiliary_DO (-0.653; 0.958), Interjection (-0.638; 2.293), Contracted_Forms (-0.585; 1.776), Analytic_Negation (-0.531; 1.435), Amplifier (-0.513; 0.907), Predicative_Adjective (-0.476; 1.083), Subject_Pronoun (-0.427; 2.219), First_Person_Pronoun (-0.302; 1.177).
		Absent features: Other_Noun (-0.669; 3.388), URL (-0.488; 4.452), Proper_Noun (-0.44; 3.477), Preposition (-0.408; 2.686), Attributive_Adjective (-0.246; 1.061), Hashtag (-0.214; 1.265), Capitalisation (-0.208; 1.089).

Other features associated with positive Dimension 2 include text-based features, such as capitalisation and colons. *Capitalisation* can be used for emphasis, as well as to compact more information into the tweet, as they frequently signal an acronym, enabling longer strings to be expressed concisely, thereby freeing up space for additional information in a character-restricted context. *Colons* are often used to introduce content and information.

Positive Dimension 2 is also characterised by CMC-specific features, such as *hashtags* and *URLs*. *URLs* allow unlimited space to extend on the content of the tweet (Yazdanfar and Thomo, 2013). *Hashtags* are often used to provide topical information (Conover et al., 2011) and broadcast the content of the tweet to an audience beyond the author's current followership, linking the content to feeds of tweets tagged in the same way without having to interact with people directly (Zappavigna, 2011). Both features are associated with an informationally dense style, as hashtags integrate topical information and URLs essentially enable tweets to include the whole contents of a website or particular webpage, despite the 280-character restriction.

Finally, *imperatives* are associated with positive Dimension 2, and these are used to direct some form of action, which are common in procedural texts (Butt et al., 2003). Overall, these features are largely indicative of an informationally dense style, comprised of numerous specific referents and carefully constructed and integrated information, characteristic of texts constructed when there is time to carefully plan and edit structures (Biber, 1988). Additionally, there is a simultaneous uni-directional broadcasting style, where content is shared to many and action is demanded from many, characteristic of non-interactive and procedural texts (Butt et al., 2003).

Whilst positive dimension 2 is characterised by features associated with a more informational communicative style, Table 15 shows that negative Dimension 2 is characterised by features indicative of an interactive style. For example, *initial mentioning* - mentioning through the '@' symbol in the initial position of the tweet – is the most strongly contributing feature, which is used to direct the tweet to another Twitter user and interact with them. *Subject pronouns*, especially *first person pronouns*, and *predicative adjectives* are strongly associated to negative Dimension 2. *First person pronouns* and *predicative adjectives* have been found to occur more frequently in interactive and involved texts, as they mark a personal focus and can be used to encode the author's stance and beliefs (Chafe, 1982; Biber, 1988; Wales, 2006). *Predicative adjectives* are also less integrated in that the structure takes up more space than attributive adjectives (e.g. *the smelly man* vs. *the man is smelly*) (Biber, 1986).

Negative Dimension 2 is also characterised by interjections and contracted forms, which are associated with informality and interactivity, as *interjections* can be used to acknowledge someone's talk and encode a reaction (Smith, 2003; Jefferson, 2002), and *contracted forms* can be used to save space or time when typing, and can be used to mirror the spoken realisation of the construction (Werry, 1996). Finally, Dimension 2 is characterised by *auxiliary do* and *analytic negation,* which often co-occur to negate a particular action or reject a previous statement, especially one that has been mentioned before (Ard, 1982), which suggests that interaction is taking place.

This opposition between an informationally dense broadcasting style with an interactive style is reflected in the tweets most strongly associated with positive and negative Dimension 2. Examples 11-15 in Table 16 are the top five tweets most strongly associated with positive Dimension 2. These tweets tend to be broadcasting information about current events, news stories and future and upcoming opportunities, often providing an exceptional amount of detail in the limited space of a tweet through complex noun phrases, integrated structures, and several nominal referents.

Table 16: The tweets most strongly contributing to positive Dimension 2

	Tweet	Coord	Contrib
11	LISTEN: John Hodgson discusses "the emperor of all conjurers" RICHARD POTTER on @nhpr: https://t.co/uJPyaiwhYy (37 minutes in). Read more about this book which the @WSJ called a "provocative record of Potter's odds-defying climb" here: https://t.co/zoAxS24sfK #magic #celebrity	0.811	0.099
12	farmdoc Webinar: @jt_hubbs and @ScottIrwinUI will review @USDA's June 29 Grain Stocks and Acreage reports and considers balance sheet and price implications for both old and new crop #corn and #soybeans. Register here: https://t.co/pqovl299qH https://t.co/dLtJZRM5HY	0.802	0.097
13	VIDEO - @odublast - "Building Leaders for Advancing Science and Technology." High school students from across the Commonwealth spent 3 days learning about to climate change, sea level rise and cybersecurity through STEM. WATCH NOW: https://t.co/YdaYTB7fJw #odu @educationODU https://t.co/DfCATdltVd	0.757	0.086
14	The Phoenix @Suns had a huge profile in yesterday's #NBADraft2018 with four picks, a big trade and brand new big man in @DeandreAyton with roots here in #Arizona. Details ahead @kjzzphoenix. Stream us here: https://t.co/qd2LqqGXs4.	0.756	0.086
15	Just signed up for WCX, the global digital currency exchange. Sign up and earn 100 X Tokens: https://t.co/83d2f2xOZj @wcxofficial #TradeTheWorld	0.747	0.084

Example 11 in Table 16 is broadcasting information on a current news story. This tweet is informationally dense, managing to include a description of the news story (*John Hodgson discusses "the emperor of all conjurers" RICHARD POTTER*), where it was discussed (*on @nhpr*), a link to the radio show, and where to find the particular section (*37 minutes in*), an instruction to read more, a review of the book they are discussing from another source (*which the @WSJ called a "provocative record of Potter's odds-defying climb"*), and another link to where interested parties can find out more, all in the space of 280 characters.

Example 12 in Table 16 is also informationally dense, managing to describe an upcoming webinar, indicating who will be presenting and what the content is going to be, and also instructing readers to sign up and register to the webinar. Example 13 is providing background information and a link to a video, describing the content of a video, and instructing readers to "WATCH NOW".

Examples 11-15 are not interacting with particular individuals, but are focused on trying to reach the broadest audience possible. Example 14 and 15, for example, are broadcast to the timelines of all their followers and also to a broader audience via hashtags, which means that anyone interested in "#NBADraft2018", "#Arizona" (Example 14) and "#TradeTheWorld" (Example 15) can also observe the tweet by searching for the hashtag.

Alternatively, Examples 16-20 in Table 17 are the tweets most strongly associated with negative Dimension 2. These tweets are considerably more interactive. The person tweeting is directly conversing with (an)other Twitter user(s), and thus, unlike the positive tweets, they have a specific addressee. Examples 16-20 include initial mentioning to direct the tweet to other Twitter users. These interactive tweets tend to form part of a Twitter conversation and are replying to something that has been said in a previous tweet, as opposed to forming the initial turn/tweet to which others can reply. For example, Example 17 is replying to a tweet sent by the McFly band member

@DougiePoynter describing how he nearly killed all his band mates in a music video. *Analytic negation* and *auxiliary do* are used in Example 17 to acknowledge what was said in @DougiePoynter's previous tweet (*I'm so glad you **didn't** killed them*). Additionally, the use of interjections "Lol" and "haha" in Example 18, and turn-initial no in "*no honestly"* in Example 20, occur to acknowledge someone's talk and mark an interactive style.

Table 17: The tweets most strongly contributing to negative Dimension 2

	Tweet	**Coord**	**Contrib**
16	@btykiwi ill spam u dont worry ! and its ok let's just support our faves tho that's why we're here but yeah i feel u 'sometimes' 😅	-0.623	0.058
17	@DougiePoynter @TomFletcher @harryjudd I'm so glad you didn't killed them 😂	-0.618	0.057
18	@TomArnold Lol haha your so stupid! Didn't you know if anything exists Mule'R has it. And guess what that means when evidence is sealed? You're going to fall... you believed him 😂	-0.616	0.057
19	@retroxirwin if you can help me that'd be amazing if not I'd understand 💕	-0.615	0.057
20	@Calliethulu no honestly, I'm sorry that must quite suck :/ *hugs if wanted*	-0.596	0.053

Many of the tweets are written as if they were spoken by including reduced forms like *contractions* and informal forms like *interjections.* The interactive tweets also tend to mark a personal focus and encode personal stance through *first person pronouns* and *predicative adjectives*, such as in Example 19 (*that'd be **amazing**),* Example 17 (***I**'m so **glad***), and Example 20 (***I**'m **sorry** that must quite suck)*.

These tweets on the negative side are also comparably shorter than the informational tweets on the positive side. Positive Dimension 2 is slightly correlated to tweet length (r = 0.21) (see section 5.1) indicating that the informational broadcasting tweets tend to be longer than the interactive tweets. This may offer support to the interpretative labels assigned to this dimension because, on the one hand, whilst it is possible for others to respond to the tweets on the positive side of the dimension, like academic writing (e.g. a journal article), there is no invitation for dialogue in these tweets, which may mean that the tweets have been designed to integrate large amounts of content. For example, in academic writing, the author compacts as much content into the article predicting particular questions and objections and acknowledging them, being as concise and efficient as possible. Although being a highly integrated style, it may nonetheless in the context of tweets take up the majority of the space afforded to them by using more words.

Interactive tweets, on the other hand, are not required to compact everything into one tweet; rather the message can be drawn out amongst several short tweets in a conversational thread. Despite being more fragmented and choosing longer structures like predicative adjectives, interaction is governed by a mechanism of exchange, which generally refers to individual participants alternating in offering relatively short bursts of information (Holler et al., 2015), suggesting that interactive turns are shorter than non-interactive turns. Thus, the slight positive correlation may be used to support the interpretation that this dimension opposes informationally dense tweets with interactive tweets.

Overall, Dimension 2 opposes an informationally dense broadcasting style with an interactive and conversational communicative style. Notably, this result reflects the two main forms of communication on Twitter: directed public conversation (one-to-one or one-to-few) and the general broadcast of information to the entire network (one-to-many) (Yaqub et al., 2017). Moreover, this opposition mirrors the three bullet points displayed on the Twitter homepage before one logs on, which are: (1) “Follow your interests”; (2) “Hear what people are talking about”, and; (3) “Join the conversation” (Twitter, 2019a). Specifically, (1) and (2) are related to the general broadcast of information, whilst (3) is associated with interactive and conversational tweets. This also confirms Pavalanathan and Eisenstein’s (2015) findings that more informal forms were used when the audience of the tweet was small, whereas more formal forms were used when the tweets contained a hashtag and therefore had a larger potential audience.

The opposition of texts oriented towards the presentation of information with texts oriented towards interaction found in this dimension of linguistic variation has been consistently identified as one of the most important dimensions of linguistic variation in numerous MDA studies of various languages and language varieties, leading Biber to suggest that this could be a universal dimension (Biber, 2014). Finding this pattern across individual tweets lends additional support to Biber’s (2014) theory. Moreover, if Biber (2014) is right that this is a universal dimension, then finding it using the modified version of MDA offers support that this short text version of MDA works. Overall, given its frequent occurrence amongst most MDA studies, this finding provides a strong basis for the interpretations of subsequent dimensions.

5.3. Dimension 3: Personal versus Other Description

The linguistic features that are strongly contributing to Dimension 3 are presented in Table 18. Table 19 presents the tweets most strongly associated with positive Dimension 3 and Table 20 presents the tweets most associated with negative Dimension 3.

Table 18: The features strongly contributing to positive and negative Dimension 3 (coordinates; contributions)

3	+	**Present features**: Other_Verb (0.243; 2.056), Past_Tense_Verb (0.276; 1.133), Infinitive (0.333; 0.984), Private_Verb (0.345; 0.91), First_Person_Pronoun (0.348; 2.781), Analytic_Negation (0.459; 1.908), Public_Verb (0.492; 1.781), Stance_Verb (0.502; 1.665), Phrasal_Verb (0.528; 1.65), Imperative (0.61; 2.657), Object_Pronoun (0.951; 6.098), Auxiliary_DO (1.034; 4.272).
		Absent features: Contracted_Forms (0.132; 0.858), Third_Person_Singular_Verb (0.179; 1.364), Predicative_Adjective (0.219; 2.417), Stative_Form (0.397; 6.405).
	-	**Present features**: Predicative_Adjective (-1.3; 14.328), Stative_Form (-0.891; 14.37), Graduation (-0.764; 2.161), Contracted_Forms (-0.707; 4.605), Demonstrative_Pronoun (-0.545; 1.547), Amplifier (-0.536; 1.762), Third_Person_Singular_Verb (-0.466; 3.543), IT (-0.422; 1.379).
		Absent features: Other_Verb (-0.356; 3.017), First_Person_Pronoun (-0.223; 1.782), Object_Pronoun (-0.123; 0.792).

Table 19: The tweets most strongly associated with positive Dimension 3

	Tweet	Coord	Contrib
21	TYLER JUST SHOWED UP AT MY WORK WITH COFFEE AND BREAKFAST FOR EVERYONE AND IF I DONT MARRY THIS MAN PLEASE SHOOT ME HOLY HELL WHAT DID I DO TO DESERVE HIM	0.672	0.12
22	@seekingBushra How about you just hop off my ass and stop bashing me for voicing my thoughts? If you don't like my tweets that much, then stop reading them.	0.651	0.113
23	Don't allow me to explore you if we are just going to end up like strangers	0.606	0.098
24	Dont wait for me to die for you to realize what you've done/lost	0.606	0.098
25	@FearlessCourt @snakelor_swift @Chriscqma @IamSeruzna @ComplexMusic Look. I don't care what y'all think, I stan someone who are using their platform to raise voice for women empowerment and lgbtq in a conservative society like South Korea's. I can stan whoever the fuck I want to, so stay pressed and salty	0.562	0.084

Table 18 shows that the most contributing feature to positive Dimension 3 is *object pronouns*, which predominantly occur in the tweets strongly associated to the dimension in the *first person* (e.g. me), which is also strongly associated to the dimension. *Object pronouns* are used to indicate that some action is acting upon or has acted upon the author or other agents, marking some form of passivity and encoding experience (Quirk et al., 1985). For example, the object pronoun occurs in Example 22 in Table 19 to describe the author's personal experience that someone has been criticising her (*stop bashing **me** for voicing my thoughts*). *First person pronouns* in object and subject position are used to involve the author into the discourse by explicitly referring to and reporting on the self (Chafe and Danielwicz, 1987), enabling one's personal stance and experience to be encoded. Examples 21-25 in Table 19 all include first person pronouns and all are used in the tweets to mark a personal focus. Overall, both kinds of pronominal features are associated with an interpersonal and involved focus (Chafe, 1982; Wales, 2006).

Positive Dimension 3 is also characterised by numerous kinds of verbs, including *stance verbs*, *private verbs*, *public verbs*, *phrasal verbs*, other more *general verbs*, *infinitives*, and *verbs in past tense* form and *auxiliary DO*. *Stance verbs* tend to be used in the tweets to express the author's desires and personal stance, such as the stance verb "want" in Example 25 in Table 19 (*I can stan whoever the fuck I **want** to*). *Private verbs* are used in the tweets, like Example 24 with "realize", to encode knowledge, thoughts and beliefs. *Public verbs* (are used to report on speech, such as 'bashing' in Example 22, and *past tense verbs* are used to narrate a particular personal past event, such as "showed up" in Example 21.

Infinitives occur in the tweets to complete the meaning of the verb and are used to expand an idea (Chafe, 1982; 1985). For example, "to explore" completes "allow" in Example 23. *Auxiliary DO* often occurs with *analytic negation* in order to negate an action, such as Example 25 (*I **don't** care what y'all think*). All of these verb forms are often used to self-report about a variety of different events, feelings, thoughts and ideas.

Finally, *imperatives* are also strongly associated to positive dimension 3 and they tend to be used in the tweets to encode the author's feelings and desires by making a personal request, such as Example 24 (*Dont wait for me to die […]*). Overall, the features strongly associated with positive Dimension 3 co-occur in the tweets in order to encode a personal experience, self-report and share what the author is doing, thinking and feeling.

By contrast, the linguistic features contributing the most to negative Dimension 3 in Table 18 are typically used for describing and evaluating a subject that is external to the self. Table 18 shows that the most contributing feature is *stative forms*, which is realised most frequently in the tweets as *BE as a main*, and is used to introduce and characterise a subject (Butt et al., 2003), often by encoding a judgement. For example, BE as a main verb occurs in Example 26 in Table 20 to encode a judgement on Dalmazzi (***It's*** *hilarious that Dalmazzi is no longer one of "THE CONSOLIDATED CASES."*).

Table 20: The tweets most strongly associated with negative Dimension 3

	Tweet	**Coord**	**Contrib**
26	@steve_vladeck Also, FWIW, it's hilarious that Dalmazzi is no longer one of "THE CONSOLIDATED CASES."	-0.657	0.115
27	@WonderBread941 That's a flaw in the system, like saying good service is perfection. 4 out of 5 stars should mean really good but keep striving to be better.	-0.609	0.099
28	@MagesticSalah @The_Koxy @Thfc_Scops @JamesPearceEcho Coutinho is so fucking good. He's a top 10 player in the world, dude is lights out.	-0.548	0.08
29	@vineet_red @ThomoJosh @SimplyUtd But Fred is one of your newest signings while Sturridge is most likely on his way out...if that's his quality then that's just underwhelming lol	-0.545	0.079
30	@bobpockrass Possible that it's not the driver but the car? Kenseth is so much better than his finishes this year.	-0.54	0.078

Negative Dimension 3 is also characterised by *predicative adjectives* and *graduation* (*comparative* and *superlative* forms), which are used for evaluating and encoding stance and descriptions of subjects (Biber, 1988; Martin and White, 2005). For instance, the predicative adjective "good" occurs in Example 28 to evaluate and describe a football player (*Coutinho is so* *fucking **good***), and the comparative form “better” is used in Example 27 and 30 to evaluate a subject in relation to something else (*Kenseth is so much **better***). *Amplifier*s, such as “so” in Examples 27 and 30 is also strongly contributing to negative Dimension 3. *Amplifiers* mark the intensity of the description and link directly to the author’s personal scale. All three linguistic features directly mark the author’s stance and enable the description of a subject.

Pronoun IT and *demonstrative pronouns* are also strongly contributing to negative Dimension 3, and are used to refer to the subject of the descriptions, especially one that is external to the self. These forms tend to indicate a shared communicative context, as the thing encapsulated in the pronoun is only deducible from the context. For example, the *demonstrative pronoun* “that” occurs in Example 27, but it is not completely clear what *that* is referring to (***That****'s a flaw in the system*). Contracted forms are also associated with negative Dimension 3, indicating that these descriptions may be informal and/or conversational. Overall, the tweets on the negative side of the dimension tend to encode the author’s opinion (*it’s hilarious…*), description (*But Fred is one of your newest signings…*), and evaluation (*He’s a top 10 player in the world, dude is lights out*) of a subject that is often external to the conversation; that is, neither the author nor (if relevant) the addressed Twitter user(s) are the subject of the tweet.

Overall Dimension 3 reflects an opposition between tweets that are personal and self-reporting with tweets that are topic-focused and characterising other entities external to the self. Like Dimension 2, this pattern of linguistic variation is not necessarily new. Grieve et al. (2010), for instance, investigated the linguistic variation in blogs and found that there was a distinction between more personal blogs with more thematic blogs. The distinction between personal and topic-focused tweets discovered in this dimension of linguistic variation of general Twitter may therefore support the common characterisation that Twitter is a micro-blogging platform because blogs similarly vary along this dimension.

In addition to reflecting blogging more generally, it could also be opposing traditional micro-blogging with more contemporary uses of Twitter. Characteristically, micro-blogging was designed for status updates (Zappavigna, 2018) - self-reports about what one is doing, thinking and feeling (Lee, 2011). Positive Dimension 3 therefore aligns with the platform's original promoted uses. Additionally, many people nowadays perceive and use Twitter as a medium for instantaneous news distribution (Sveningsson, 2014; Zappavigna, 2018). Characterising external entities is important in the process of describing and disseminating newsworthy information. This dimension may therefore be opposing traditional micro-blogging with one of Twitter's modern day uses - for disseminating news.

Although this may not be especially surprising, what is interesting is that this dimension finds a distinction between stative forms with all other verbs. This dimension indicates that a variety of verbs are frequently used when referring to the self, whereas only stative forms tend to be used when describing others. In a study on the spoken and written language of children and adolescents, McGuire and McGuire (1986) found a similar pattern, where their participants used a wider variety of verbs when referring to the self, whilst only stative verbs tended to be used when describing other people.

They suggest that their findings indicate that the self is thought of concretely in terms of what one does, whereas other people are thought of more abstractly in terms of what they are. They argue that this is because of the experience and access that one has with the self in a variety of situations, which means that the self can be described with respect to what they are doing, thinking and feeling more easily than when describing other entities or people. Other people or entities are witnessed in a smaller number of situations, and therefore can only be thought of in terms of general dispositions.

5.4. Dimension 4: Promotional versus Oppositional

The linguistic features that are strongly contributing to Dimension 4 are presented in Table 21. Table 22 presents the tweets most associated with positive Dimension 4. Table 21 shows that positive Dimension 4 is characterised by CMC and Twitter-specific features like *hashtags, URLs* and *non-initial mentioning*, which can be promotional resources of tweets (Page, 2012). *Hashtags* are promotional in that they are often used to gain visibility and promote the content of the message to an audience beyond one's followers (Page, 2012). For example, in Example 35 in Table 22, the hashtag "#BTSxLotteFamilyConcert", concerning the concert put on by the k-pop band BTS (Bangtan Sonyeondan) will likely have been used and accessed by a global audience, especially those unable to attend the concert personally, as Twitter users follow certain hashtags, which align with their interests.

Table 21: The linguistic features strongly contributing to positive and negative Dimension 4 (coordinate; contribution)

4	+	**Present features**: Capitalisation (0.242; 0.835), URL (0.284; 2.09), Time_Adverb (0.422; 0.99), Exclamation_Mark (0.455; 1.896), Predicative_Adjective (0.464; 1.858), Object_Pronoun (0.47; 1.513), Hashtag (0.476; 2.172), Amplifier (0.501; 1.565), Emoticon/Emoji (0.504; 3.028), Subject_Pronoun (0.545; 6.515), Stance_Verb (0.586; 2.301), Perception_Verb (0.587; 1.407), First_Person_Pronoun (0.663; 10.24), Non-Initial_Mention (0.757; 3.523), Contracted_Forms (0.787; 5.811).
		Absent features: Attributive_Adjective (0.164; 0.854), Initial_Mention (0.269; 2.765).
	-	**Present features**: Question_Mark (-0.9; 4.755), WH-Word (-0.688; 3.015), Auxiliary_DO (-0.62; 1.563), Passive (-0.594; 1.082), Initial_Mention (-0.478; 4.91), Nominalisation (-0.462; 1.544), Analytic_Negation (-0.394; 1.427), Definite_Article (-0.307; 1.566), Third_Person_Singular_Verb (-0.269; 1.2), Attributive_Adjective (-0.185; 0.959).
		Absent features: First_Person_Pronoun (-0.425; 6.561), Subject_Pronoun (-0.317; 3.793), URL (-0.218; 1.601), Contracted_Forms (-0.147; 1.082).

URLs can be promotional, as they are used to promote and share additional content, such as Example 33, which is a link to attract viewers to a website and live stream of the author, who is an online gamer. Additionally, *URLs* can support the content of the tweet, such as Example 32, which is an image of the author with a box of cereal (*thank goodness for these crunchy flakes!*), and Example 35, which is an image of a BTS band member.

Non-initial mentioning can also be a promotional resource, as it is often used in the tweets to refer to other Twitter users, as a way to promote their feeds to others, as in Example 33 (*check out these peeps! -* ***@AyrockMusic @Alyssa_Rxse @chasifrass4*** *[...]*). Moreover, rather than interact with the Twitter user via initial mentioning, which directs the tweet to the particular user and reduces the visibility of the tweet to only those who follow both the author and the mentioned user, *non-initial mentioning* can be used to increase visibility, as the tweet is posted to the mentioned Twitter user *and* the Timelines of all of one's followers, increasing the chance of it being noticed. All of these features are used to promote and increase visibility, not only of the content of the tweet, images and other additional multimedia content, but also to promote other Twitter users and events.

Table 22: The tweets most strongly associated with positive Dimension 4

	Tweet	Coord	Contrib
31	TGIF #itscalledfashionsweatylookitup and I'll be in all day tomorrow too - come see us if you want this di*khead lewk 🍄✨ @ Claw and Co. https://t.co/DJ82bMx2Ht	0.673	0.123
32	Breakfast/lunch time! I'm feeling really sick and this is all I could stomach so thank goodness for these crunchy flakes! Now send me to @LoveIsland please! #kelloggscerealdater #hungry #crispyflakes https://t.co/zW9L6kyDKX	0.667	0.121
33	IM LIVE! SO PEEPS COME SAY HELLUR! ❤️🙏🏽 https://t.co/uKt2TTMHxU Also check out these peeps! - @AyrockMusic @Alyssa_Rxse @chasifrass4 @PROJ3CTGD @_OnlyVic_ @MrSynnn @Joshh9761 @xLeaahh @OrbzTV @KingTylerish @lolgabiz @Elthirlwell	0.655	0.116
34	Ugghhhh!!! I'm literally screaming!!!! Jungkook really looks so good and hot! 😭💜❤️💗❤️💖💋💕💖💓😍#JUNGKOOK#RedHairJungkook @BTS_twt	0.643	0.112
35	I've been wishing for a comeback of Kim Seokjin in blonde since last year.. and now here it is 😭 my heart is so happy, im so happy 😭💜 @BTS_twt thank youso much for this 😭🙌🏻 #BTSxLotteFamilyConcert https://t.co/w7OnQBSAqv	0.616	0.103

Table 21 also shows that positive Dimension 4 is characterised by *first person pronouns, stance verbs, perception verbs, predicative adjectives* and *emojis,* all of which are associated with encoding personal stance. Examples 31-35 all mark a personal focus through first person pronouns. Additionally, stance verbs and predicative adjectives are used in the tweets to explicitly encode stance, such as in Example 35 (*I've been* ***wishing*** *for a comeback of Kim Seokjin… my heart is so* ***happy***). Perception verbs mark a sensory description, enabling descriptions of a particular viewpoint, (*Jungkook really* ***looks*** *so good and hot*!) and particular states (*I'm* ***feeling*** *really sick*), and emojis are often used to encode emotional states and support the sentiment expressed in the tweet. For instance, the emoji with heart eyes used in Example 34 is a way of expressing that something is attractive, which supports the description that Jungkook looks "hot".

Table 21 also shows that positive Dimension 4 is characterised by *amplifiers*, *exclamation marks*, and *capitalisation*, which are associated with an emphatic production and suggest a confident and excitable conviction, such as Examples 33 and 34. Finally, time adverbs are also associated with positive Dimension 4, and they tend to be used to provide temporal information in situational contexts where time is relevant and when the audience is not physically present (Bohmann, 2017), such as to promote what time an event is going to happen, like in Example 31 (*I'll be in all day* ***tomorrow*** *too - come see us…)*.

Overall, when viewed in the context of the tweets (see Table 22), these features are connected by an underlying promotional communicative function. Specifically, the tweets are often making endorsements, whereby the author encodes their stance and expresses approval (*I'm literally screaming!!!!!*), support (*BLONDE JIN IS SUPERIOR!!!*), and gratitude (e.g. ***thank goodness*** *for these crunchy flakes […], TGIF* (Thank God It's Friday)) for the entity they are endorsing. Moreover, the tweets on the positive side often promote entities (self or other) by incorporating recommendations (*come see us if you want this di*khead lewk*).

Alternatively, the features most strongly associated with negative Dimension 4 in Table 21 are used in the tweets, presented in Table 23, in order to oppose a particular stance or entity, as opposed to endorsing and promoting it. For example, negative Dimension 4 is characterised by features associated with interactivity and asking questions, such as *initial mentioning, WH-words* and *questions marks*, which are used in the tweets most strongly associated with negative Dimension 4 to introduce a counter-argument and challenge a previous statement from another Twitter user in the form of a question, such as Example 38 (***What*** *do you think PlayBoy Magazine is where almost most famous people pose nude**?**).* Questions are also used in the tweets to express disbelief in order to ridicule and oppose the particular position, such as Example 37 (***What*** *are you talking about**?**).* A present tense orientation is marked through these questions features and *third person singular verb forms,* which are often used when the topic is of immediate relevance (Biber, 1988), characteristic when opposing and providing counter-arguments, as the topic becomes relevant and needs to be opposed.

Negative Dimension 4 is also characterised by *analytic negation* and *auxiliary DO,* which are used to negate an action so as to oppose it, such as Example 40 (*honestly* ***don't*** *understand why people eat animal products*), and

Example 37 (*No, this country was**n't** founded by immigration*). Additionally, *auxiliary DO* often occurs in the tweets to emphasise the verb and indicate a counter view. For example, in Example 38, *auxiliary DO* is used to mark disagreement with a previous statement: “But all models **do** pose somewhere nude.” Whilst the tweet prior to this tweet is unavailable, it can be assumed that the previous tweet suggested that not all models pose nude.

Table 23: The tweets most strongly associated with negative Dimension 4

	Tweet	Coord	Contrib
36	@zazingwa @CryptoPseudonym @Crypto_Trogdor @cryptostardust @Elbrusco3 @mrdmrts @cripdohsimpson 4/ Can you explain why anyone would pay the astronomical price to fake the launch of hundreds of satellites, and hundreds of thousands of similar videos/images? What benefit is gained by such an extraordinary, and arbitrary, conspiracy?	-0.626	0.106
37	@iamkhaledd No, this country wasn't founded by immigration... What are you talking about?The British came here and conquered the land, then maintained it.	-0.619	0.104
38	@TanaLara4 @Punkster1011 But all models do pose somewhere nude. What do you think PlayBoy Magazine is where almost most famous people pose nude? Isn't that "Adult" magazine?	-0.57	0.088
39	@adjani_98 Right? How much common sense does it take understand how they can't make more money by helping less people have it....	-0.565	0.086
40	honestly don’t understand why people eat animal products when plant-based tastes better, is healthier, has less of an environmental impact and doesn’t hurt animals. Kind of a no-brainer, don’t you think?	-0.548	0.081

Negative Dimension 4 is also characterised by features indicative of a detached style and high abstract informational focus, such as *passive constructions*, *nominalisations* and *definite articles* (Chafe and Danielwicz, 1987; Biber, 1988). Detached language involves subjects that are abstractions (Chafe and Danielwicz, 1987). For instance, “benefit” in Example 36 (*What **benefit is gained** by such an extraordinary, and arbitrary, conspiracy?*). *Passive constructions* also enable the thing or entity affected by the action to be emphasised, often an abstraction, which occurs in the tweets to provide a counter argument. In Example 36, the ‘conspiracy maker’ is not explicitly mentioned in the sentence, but instead the benefit and beneficiary of the conspiracy is emphasised, as a way to suggest that there is no beneficiary in order to oppose the addressee and imply that they are wrong. As opposed to positive Dimension 4, the tweets on the negative side of Dimension 4 in Table 23 are considerably less supportive and instead function to oppose, challenge, and dispute a particular stance, statement or ideology.

Overall, Dimension 4 reveals a distinction between tweets that are promotional with tweets that are non-promotional and oppositional. As the fourth major dimension of linguistic variation, it indicates that promoting oneself and others and opposing others are important and common communicative tasks on Twitter. The promotional function reflected in the linguistic features and tweets most associated with positive Dimension 4 supports Page (2012), who has suggested that practices of self-branding and micro-celebrity are very frequent on Twitter.

The promotional side, however, is not only comprised of self-promotional tweets, but they are also promotional in that they endorse and recommend particular entities, including other Twitter users. One of Twitter’s major uses is for following celebrities, businesses, organisations and people/groups/topics that users are interested in, as well as communicate with them. This practice of promoting and showing support for other entities is arguably associated with demonstrating affiliation and fandom (Zappavigna, 2018), and following one's interests, which aligns with one of the major promoted uses of Twitter.

Twitter and social media more generally has been described as fostering echo-chambers, as individuals get to choose what content they see, meaning that the users are more likely to choose to see content that supports their views (Krasodomski-Jones, 2016), called ideological homophily. The oppositional side of this dimension of linguistic variation suggests that oppositional content is rife on Twitter, indicating that many individuals do seek out or choose to encounter oppositional content. This supports Vaccari et al. (2016), who found in their study investigating political expression that political homophily was not a universal outcome and that many Twitter users encountered oppositional content.

5.5. Dimension 5: Persuasive versus Non-persuasive

The linguistic features that are strongly contributing to Dimension 5 are presented in Table 24. Table 25 presents the tweets most associated with positive Dimension 5 and Table 26 presents the tweets most associated with negative Dimension 5.

Positive Dimension 5 is characterised by numerous features associated with an interpersonal function and a shared communicative context, such as *imperatives, WH-words, question marks, second person* *pronouns, interjections, demonstrative pronouns, private verbs* and *complementation*. Additionally, *capitalisation* and *exclamation marks* are also strongly associated, which can be used for emphasis (Smith, 2003) or to mimic oral production and denote shouting (Postmes et al., 2000).

Table 24: The linguistic features most strongly contributing to positive and negative Dimension 5

5	+	**Present features**: Non-Initial_Mention (0.357; 0.91), Private_Verb (0.389; 1.366), Interjection (0.444; 2.334), Demonstrative_Pronoun (0.454; 1.271), Capitalisation (0.459; 3.497), Auxiliary_DO (0.528; 1.317), Hashtag (0.606; 4.098), Second_Person_Pronoun (0.735; 8.679), WH-Word (0.759; 4.272), Complementation (0.818; 2.827), Question_Mark (0.971; 6.432), Exclamation_Mark (0.972; 10.085), Imperative (1.127; 10.735).
		Absent features: Past_Tense_Verb (0.148; 1.133), First_Person_Pronoun (0.152; 0.982), Subject_Pronoun (0.157; 1.083), Other_Noun (0.418; 2.78).
	-	**Present features**: Passive (-0.802; 2.29), Profanity (-0.55; 1.314), Past_Tense_Verb (-0.437; 3.352), Third_Person_Pronoun (-0.379; 1.552), Indefinite_Article (-0.306; 1.414), Subject_Pronoun (-0.27; 1.86), First_Person_Pronoun (-0.238; 1.532), Other_Noun (-0.124; 0.827).
		Absent features: Second_Person_Pronoun (-0.221; 2.609), Exclamation_Mark (-0.176; 1.829), Imperative (-0.156; 1.487), Capitalisation (-0.144; 1.098).

Whilst many of these features are associated with interactivity and a shared communicative context, positive Dimension 5 is characterised by *hashtags* and *non-initial mentioning*, which are used to broadcast the tweet to particular feeds and refer to other Twitter users in the third person, as opposed to interacting with them directly. The occurrence of these features amongst interpersonal features suggests that these tweets are not necessarily interacting with specific people directly, but are perhaps purposely employing an interpersonal style and presupposing a shared communicative context in order to imply friendliness and intimacy, and perhaps address a particular, albeit unspecified audience.

Table 25: The tweets most strongly associated with positive Dimension 5

	Tweet	Coord	Contrib
81	ARE YOU HUMAN TOO? 💜 Where did you find that extra charm? You're killing me softly! WHY SO HOT?? PLEASE EXPLAIN 😄😍 @BTS_twt @bts_bighit #BTSxLotteFamilyConcert https://t.co/oUhE8dkYZy	0.702	0.156
82	Find out how much you could save switching to BrandStencil for your artwork creation - use our ROI calculator! https://t.co/vro5K44Q2n #MarketingAutomation	0.644	0.131
83	Attn, #mixers! Stop what you're doing and listen to @LittleMix and @CheatCodesMusic's #OnlyYou 🔥🔥🔥†#littlemixONLYYOU https://t.co/AjkcT0zWqV	0.643	0.13
84	All #robots need drive systems to move around but do you know how it works? Join our #STEM courses for free! https://t.co/RCmUjar9kX	0.641	0.129
85	Where are our #DogsofHOS at? Don’t forget to tag us in your photos so we can get over-excited and share them everywhere! 🐶 https://t.co/tBZEzn1tZn	0.627	0.124

By looking at the tweets most strongly associated to positive Dimension 5 (see Table 25), it is clear that these features are all connected by an underlying persuasive communicative function. Specifically, the tweets aim to bring about something, often employing *imperatives* to demand some action. For example, Example 41 demands an explanation (*PLEASE EXPLAIN*) and Example 43 instructs people to “Stop what you're doing and listen to […]” a new music track.

Although *imperatives* direct future action, they are not always persuasive. However, certain kinds of *imperatives* can contain presuppositions as a way to persuade the reader. For example, in Example 42 the *imperative* presupposes that the reader is overpaying for their artwork creation, and suggests that they could save money by switching to their particular company (*Find out how much you could save switching to BrandStencil […]*).

Additionally, many of the conversational and interpersonal features described previously are used in the tweets most associated to the dimension to persuade. For example, *question marks, WH-words* and *auxiliary DO* are used in the tweets as instances of questions, which often encourage particular responses, as a way to persuade the reader to do something. Example 44 has the question “but **do** you know how it works**?**”, which encourages the response “no”, and by extension problematises a lack of knowledge. Nevertheless, a solution is offered in the imperative clause “Join our #STEM courses for free! https://t.co/RCmUjar9kX”, which persuades the reader to sign up and join in order to find out.

Moreover, *second person pronouns* often occur presuppositionally in the sense that the addressee is assumed to read the tweet, as opposed to the author targeting the tweet directly to them through initial mentioning. Second person pronoun occurs in Examples 42, 43 and 44 in order to target readers directly without mentioning them specifically. Such a technique is used commonly in advertising language (Cook, 2001). Overall, an interpersonal and informal style characteristic of many of the features associated to positive Dimension 5 have been described as effective for persuading in advertising discourse as they reduce the distance between the addressee and imply friendliness and intimacy (Wells, Moriarty, and Burnett, 2006; Cui and Zhao, 2013), which may persuade the reader to trust what is being said.

Overall, the tweets associated with positive Dimension 5 are persuasive, often trying to bring about something in the future and promoting particular products, companies, and content, using various strategies to persuade the reader to engage and invest with them.

Alternatively, the linguistic features strongly associated to negative Dimension 5 are far less associated with persuading the reader to do something, rather the features are indicative of personal narratives. Specifically, *past tense verbs* and *passive constructions* are the most strongly associated features with this side of the dimension, and they tend to occur in the tweets to refer to past events and experiences. For example, past tense verbs and passive construction occur in Example 46 in Table 26 to report that the author's "ACCOUNT JUST **GOT** TEMPORARILY **LOCKED**". Additionally, passive construction occurs in Example 48 to report that "Even Womble the Therapy Dog **gets involved**" in the team development days.

Subject pronouns are also strongly associated with negative Dimension 5, especially *first person* and *third person personal pronouns*, and these are used to mark personal narratives, such as Example 46 (***I** JUST HAD A MINI HEART ATTACK*), as well as narratives about external entities, such as Example 47 (***I** had **him** undress all the way*). *Indefinite articles* and *general nouns* are also associated to negative Dimension 5, which are indicative of new content being presented and described, and are used in the tweets to provide detail on the past experience, such as Example 50 (*was meant to be **a surprise***).

Table 26: The tweets most strongly associated with negative Dimension 5

	Tweet	Coord	Contrib
46	bITCH MY ACCOUNT JUST GOT TEMPORARILY LOCKED. I JUST HAD A MINI HEART ATTACK	-0.455	0.065
47	Free watching of young male gay sex I had him undress all the way https://t.co/ynr3m3YHPc telugu hot gay porn zeb atlas fuck twink videogaysex afganistangayporn gay straight bait bus video full 3gp chloroform muscled gay studs voyeur gay medic	-0.43	0.058
48	We love our team development days at Lighting up Learning once every six weeks. Such unified and synchronous thinkers whilst deep respect enables them to challenge each other, including our Director. Even Womble the Therapy Dog gets involved. https://t.co/5KoNLSxUes	-0.427	0.057
49	Octavia E. Butler's 'Kindred' remains one of the best books I have ever read. So good I had no idea it was published in 1979 - timeless.	-0.411	0.053
50	Update: was meant to be a surprise but just realised my location is still on so she probably already knows I'm here. Fucked it	-0.41	0.053

The final feature strongly associated to Dimension 5 is *profanity*. Whilst profanity can have several functions, it is used in the tweets mainly to evaluate the event, often as an amplifier or immediate response, such as in Example 46, or incorporated into the description of the event, such as Example 47 (***fuck*** *twink videogaysex*), or as an overall analysis, such as Example 50 (***Fucked*** *it*), suggesting a personal focus.

Overall, tweets on the negative side of the dimension, presented in Table 26 are considerably less persuasive, but instead are often narrating personal past events and experiences. Additionally, these tweets often contain personal judgements and reflections, such as Example 49 (*Octavia E. Butler's 'Kindred' remains one of the best books I have ever read*).

Dimension 5 finds a distinction between tweets that are functioning to persuade and bring something about in the future with tweets that are functioning to narrate and evaluate past events. This suggests that Twitter is used more generally to obtain some form of capital (economic or non-economic), as other Twitter users are persuaded to do something.

Biber (1988) found a dimension of functional linguistic variation across spoken and written English, which was assigned the label ‘Overt expression of persuasion’, and consisted of the features *infinitives, modals of prediction, suasive verbs, conditional subordinators, necessity modals, and split auxiliaries*. Interestingly, most of these features do not occur in more than five percent of the tweets and have either been removed (e.g. necessity modals), or pooled with other broader categories (e.g. conditional subordinators -> general subordinators, suasive verbs -> general verbs). Despite these broader categories not being in the list of features with strong contributions, some of the features occur in the tweets strongly associated with the persuasive side of the dimension. For example, suasive verbs (e.g. move) are present in the examples. Additionally, infinitives and modals of prediction occur in the feature set in this study and these are both associated with the persuasive side of the dimension (although not exceptionally strongly), corresponding in part with Biber’s (1988) findings. Moreover, in Biber’s (1988) research, infinitives were realised as complement clauses, and in the present dissertation *complementation* is strongly associated with the persuasive side of the dimension. Importantly, Biber (1988) did not have the features *imperatives, interjection*s or *auxiliary DO* in his research nor were there specific CMC features such as *non-initial mentioning* and *hashtags*. Some of these additional features contributing to a persuasive function in the present study may also occur frequently in the texts strongly associated with Biber’s persuasive dimension. Nevertheless, these unique features illustrate how persuasion is frequently done on Twitter.

In addition to persuasion, this dimension also reveals the communicative function of narrating personal past events. Another one of Biber’s (1988) dimensions that has also been found across numerous MDA studies is the narrative communicative purpose, which include features such as past tense verbs and third person pronouns, like the negative side of Dimension 5 in this analysis. Whilst there is no doubt that the features are narrative, this dimension also includes first person pronouns, which indicate that these narratives are more personal. Biber and Egbert (2016) observed a similar pattern in their MDA of the searchable web, although they interpreted it as an oral narrative dimension.

Titak and Roberson (2013) also found a dimension of functional linguistic variation in their study on web registers that opposes texts that have a past orientation with texts that have a present orientation. Many of the features associated with their past orientation are observed in the features strongly associated with negative Dimension 5 in this chapter. For example, past tense verbs and subject pronouns like third person pronouns and first person pronouns. Using Titak and Roberson’s (2013) dimensions, Friginal, Waugh and Titak (2018) found that Tweets were more associated to the present orientation than the past, albeit political tweets. These findings are at variance with the findings in this research. Although Twitter (2019b) encourages users to “See what’s happening in the world right now”, which arguably is suggestive of a more present tense orientation, this dimension reveals that a personal narrative function is also a major pattern of linguistic variation on Twitter, and thus a major communicative purpose of Twitter users is to narrate past events.

5.6. Discussion

This dissertation aims to provide a thorough linguistic description of Twitter trolling, not only with respect to its major communicative functions and patterns of linguistic variation, but also with respect to the major communicative functions and patterns of linguistic variation of general Twitter. However, in addition to Twitter trolling, there had not been an investigation into the major patterns of linguistic variation found across general English Twitter. Consequently, this chapter sought to fill this gap by identifying and describing the major patterns of linguistic variation across general English Twitter by applying a short text version of MDA to a corpus of general English tweets. After controlling for text length in the first dimension, this chapter identified 4 subsequent dimensions of linguistic variation of general English tweets (see Table 27), which were interpreted for their underlying communicative function.

The second dimension opposes tweets that are informational broadcasts with tweets that are interactive. This major dimension has been observed in almost all studies employing MDA, and thus offers support to Biber’s (2014) hypothesis that this is a universal dimension. Simultaneously, it reflects the two main communicative patterns found on Twitter, which is directed one-to-one conversation and the public broadcast of content (Yaqub et al., 2017).

Table 27: Summary of Dimensions of Linguistic Variation of General English Twitter Corpus

Dim		Summary
1	Long tweets vs. Short tweets	• This dimension opposes the presence of features with the absence of all features. • Tweets' Dimension 1 coordinates are strongly positively correlated to text length. • The length of tweets is the strongest influence on the presence of features – the more words a tweet has the more likely it will have the presence of features. • Tweet length was not controlled for in this short text version of MDA when analysing the presence or absence of linguistic features, as opposed to their relative frequencies.
2	Informational broadcast vs. Interactive	• This dimension opposes many noun types, noun modifiers, URLs and hashtags, associated with careful integration and broadcast of information with initial mentioning, pronouns, stance related features, and features associated with a shared communicative context, characteristic of interaction. • Corresponds with Biber's (1988) first dimension and most other MDA studies (Biber, 2014). • Aligns with the two major communicative patters found on Twitter (one-to-many broadcast versus one-to-one/few public interaction). • Conforms to the promoted uses of Twitter.
3	Personal vs. Other Description	• This dimension opposes tweets that are highly personal and report on the self about what one is doing thinking or feeling with tweets that are characterising and describing entities external to the self. • Corresponds to previous research on blogging, which found an opposition between more personal blogs with topic-focused blogs (Grieve et al. 2010). • Aligns with the traditional purpose of micro-blogging – to report on the self. • Is consistent with McGuire and McGuire (1986) on the opposition between numerous verb types when referring to the self, whereas only *BE as a main verb* is used to characterise others.
4	Promotional vs. Oppositional	• This dimension opposes tweets that are showing support for and promoting the self or other entities by gaining visibility with tweets that are opposing and challenging a particular statement. • Promotional style and practices of micro-celebrity and self-branding, and oppositional style may be a result of importance placed on follower size (Page, 2012; Anderson, 2019; Dorsey, 2019). • Oppositional style may be a result of the prevalence of trolling and diversity of people and opinions of Twitter.

5	Persuasive vs. Non-persuasive	• This dimension opposes tweets that are persuading other Twitter users to do something and bring about something in the future with tweets that are non-persuasive but are reporting on past events. • Suggests that Twitter is used for some form economic or non-economic gain. • Suggests that a narrative communicative purpose is important on Twitter. • Partially aligns with Biber's (1988) narrative dimension, Biber and Egbert's (2016) oral narrative dimension, and Titak and Roberson's (2013) past tense orientation dimension. • May be explained partially by the high frequency of businesses and companies using Twitter.

The third dimension finds an opposition between tweets that are personal, often reporting on the self with tweets focused on describing other entities. This dimension of linguistic variation has been interpreted as corresponding to the common characterisation of Twitter as a microblogging service because a similar opposition was observed in Grieve et al.'s (2010) research on blogging, which found a distinction between blogs that are personal with blogs that are topic-focused.

The fourth dimension contrasts tweets functioning to promote and endorse particular entities with tweets functioning to oppose. Unlike the second and third dimension, which are patterns previously found in MDA research, this dimension identifies a specific yet very important function of Twitter – as a platform for gaining attention and positioning oneself. This dimension of linguistic variation may be influenced by following one's interests on Twitter, where individuals may be more inclined to express support towards and promote one's interests, and may at the same time come into contact with interests at variance to theirs and feel the need to express opposition. For example, example 40 is on the topic of food; specifically it is pro-vegetarianism/veganism and anti-meat eating. Such a topic will spark interest from people within both camps: pro-meat eating and pro-vegetarianism/veganism, which means that the likelihood of observing an opinion that is at variance to one's own is increased. When Twitter users observe alternative viewpoints, they may be provoked to provide counter-arguments and or express opposition to these statements. Thus, tweets expressing opposition may be a result of people with oppositional interests and likes and dislikes.

Another possible reason for the oppositional communicative function may be due to the pervasiveness of trolling on Twitter. One strategy of trolls has been to take the opposing view even if they do not believe it themselves in order to provoke a response (Phillips, 2016). Phillips (2011), for instance, describes how trolls need to be emotionally detached, as they exploit and oppose the emotional sensitivities of others in order to provoke a response. At the same time, if trolls manage to provoke a reaction, the victim's reaction can be oppositional too, as users play the "I'm right, you're wrong and my opinion is better than yours" game. Thus, the oppositional style reflected in this dimension may be a result of trolling and the interaction that takes place between trolls and their victims.

The high degree of promotional tweets may be because Twitter is used by various individuals, including celebrities, as well as businesses and organisations, whose main purpose is to promote their brand, activity and products. For example, celebrities use Twitter to promote their activities and engage with their fans. Twitter also accommodates to the need for businesses to promote their brand and products as it provides a function for businesses and individuals to promote their tweets and direct advertisements to particular groups (at a fee) called 'promoted tweets'.

Another possible explanation for this dimension pattern may be due to the architectural characteristics of Twitter. Importantly, this dimension opposes promotional tweets with oppositional tweets, both of which have been described as Twitter practices that have been influenced by the importance placed on follower size. A major and important feature of Twitter is the non-reciprocity of following practices (see section 2.2.1), meaning that a person can have millions of followers and tweet and broadcast their messages to them without having to subscribe to their follower’s updates. The size of a follower list is often taken as a sign of status (Page, 2012). Jack Dorsey (2019), the CEO of Twitter, recently discussed in a TED interview how follower size is incentivised via the architecture of Twitter because the number of followers that a user has is presented in big and bold on one’s profile. It has been argued that the importance placed on follower size as a result of these architectural decisions in the beginning has influenced certain Twitter practices, such as promoting, self-branding, micro-celebrity (Page, 2012), and oppositional, provocative and uncivil content (Anderson, 2019). Specifically, it has been shown that being provocative can lead to an increase in attention and an increase in followers (Anderson, 2019). Moreover, previous research has found that such oppositional tweets get more retweets (Stieglitz and Dang-Xuan, 2013), which similarly incentivises more oppositional content, as one is more likely to be noticed. Overall, the need for more followers is offered as a possible explanation for both of the communicative functions of tweets opposed in this dimension, and this need is a result of the importance placed on the size of followers - architecturally and semiotically. Essentially, it can be argued because follower size was enlarged and put in bold, follower size is regarded as a sign of status, and thus gaining followers has become an incentive for Twitter users. This incentive has therefore influenced the high frequency of promotional and oppositional tweets, which are important communicative styles for gaining attention.

Finally, the fifth dimension of linguistic variation reveals an opposition between tweets functioning to persuade and bring about something in the future with tweets with a past tense orientation. This dimension of linguistic variation suggests that Twitter is used for economic or non-economic gain, as a persuasive communicative style is used to obtain something from others and bring about something. Tweets that are persuasive incorporate a range of different text types, including advertisements, political tweets, as well as tweets which mark the author’s point of view. In recent years, Twitter has become a platform for political expression, not only by politicians themselves, but also ordinary citizens. Rhetoric and persuasion is expected in political tweets, whereby the ideologies of one party are expressed in order to influence particular outcomes.

Moreover, Twitter is a commercial enterprise, which receives at least 86% of its revenue from advertisers (Beers, 2019) by selling the ability to promote products, tweets, accounts and trends to consumers. Advertisers importantly pay Twitter to broadcast content to the right audience, which persuades consumers and promotes brands. Promoted tweets are seen by most, if not all users; however, the kinds of promoted tweets users experience will vary from one user to the next, as they are designed and tailored to particular identities using algorithms so that adverts reach the right audience. Because promoted tweets, regardless of the content, are seen by most users, their communicative style is observed and may set a precedent for users as to 'how to tweet'. Thus, not only does the purpose of Twitter as a commercial enterprise influence a high degree of promoted tweets, which are often characterised by a persuasive (positive Dimension 5) communicative style, but because promoted tweets are observed by most users, their style may be used as a linguistic frame.

At the same time, people often use Twitter to follow their interests, which may mean that they are likely to come into contact with points of view that are at variance to theirs. They may therefore attempt to persuade those individuals to think otherwise, especially if they are particularly passionate about the subject, by encoding their personal viewpoint.

One possible explanation for a high frequency of tweets that have a narrative communicative style could be that news reportage is often shared on Twitter with great ease, and news reportage has a past tense orientation. For example, online news sites provide a small Twitter logo on the page of particular articles, which can be clicked in order to share the story with one's followers and add commentary. Moreover, Twitter has a 'Moments' page, where past events are reported on and described. Finally, many of the tweets strongly associated with this side of the dimension are characteristic of status updates, where personal recent events are reported on and described. Posting tweets is also referred to as updating one's status. Therefore, one explanation for the high degree of past tense orientation may be the result of the notion of 'update', which demands bringing one's readers up to speed by talking about not only what is happening, but also what has happened since the last update.

The results of the analysis suggest that there is a great deal of linguistic variation on Twitter. Like Honeycutt and Herring (2009) and Yaqub et al. (2017), the results show that interaction and conversation, and the public broadcast of information are important communicative goals of tweets. Additionally, like Sardinha (2018), who found that stance-taking was important across Internet registers, the results of this chapter suggest that stance-taking is important across Twitter, as we encode stance when we report on our thoughts and feelings, when we characterise and describe external entities, when we show support and promote particular things, when we express opposition, and when we persuade people to do something or think in a particular way. Stance-taking is a means for identity construction (Bucholtz and Hall, 2005), as by encoding our stance on particular topics, we position ourselves and construct an identity that aligns or disaligns with them.

The results show that stance-taking on Twitter varies considerably as it intersects many of the major linguistic repertoires, which suggests that it is important for a variety of different purposes across Twitter, including identity construction. Finally, the range of stylistic variation found across Twitter in this analysis is indicative that Twitter is used commercially for social and economic gain (i.e. profit-making, including gaining money, attention, support, followers, likes, retweets, social interaction), suggesting that users of Twitter have common communicative goals, which exploit the technology's potential of reaching particular audiences.

5.7. Conclusion

On Twitter, and social media generally, individuals are provided with the opportunity to manage how they present themselves. Each person creates a profile and they are able to present themselves and construct their identity through their profiles and messages. Whilst this “is in many ways liberating”, it has also “brought pressure to create compelling identities that attract attention” (Baym, 2014: 222), and in the case of Twitter, attract followers. This suggests that many people view their profiles and their messages as something to be consumed and that there is an underlying need to gain attention and gain more followers. The range of stylistic variation found across Twitter supports this, showing that Twitter is commonly used for self-commodification, as people manage their identities, engaging in practices of self-branding through stance-taking, self-reporting, promotion and persuasion, as well as broadcasting their message beyond their followership, distributing news, and expressing opposition and this often occurs in order to attract attention and gain something. Additionally, the results show that interaction is important, suggesting that Twitter is also used for social and interpersonal gain.

Despite using a modified version of MDA for this analysis, the results of the MCA have been comprehensible and clear. In particular, MCA identified the dominant relationships between the tweets and the categories of linguistic features, and the major patterns of variation amongst them. This chapter has shown that these relationships and patterns of variation are a result of an underlying communicative function, offering support to the notion that linguistic features vary because they serve communicative functions (Brown and Fraser, 1979). Overall, these results have been satisfying with respect to their coherence and in demonstrating the utility of short text MDA in identifying the major patterns of linguistic variation across a corpus of short texts. This importantly broadens MDA's reach to the analysis of linguistic variation between individual short texts, as up until now, it has tended to be applied to longer texts or short texts that have been concatenated (see section 4.6).

With respect to the major aim of this dissertation, the results of this analysis indicate the severity of oppositional content on Twitter. This finding perhaps runs counter to findings that Twitter fosters homophily and echo-chambers (Yang and Eisenstein, 2017; Balusu, Merghani and Eisenstein, 2018; Krasodomski-Jones, 2016), although the degree to which people seek out alternative positions is not known. Whilst access to opposing views and opinions can be beneficial and important, it can, at the same time turn sour very quickly. As the fourth major dimension of linguistic variation, it suggests that oppositional and challenging content is rife across Twitter, and this may be a result of the prevalence of trolling and trolls who purposely take the opposing view to provoke a reaction. Whilst previous research has clearly demonstrated that trolls can be oppositional (Phillips, 2016), the range of linguistic variation across Twitter trolling has yet to be examined. The following chapter presents the results of the short text version of MDA on the corpus of Twitter trolling.

6. Dimensions of Twitter Trolling

To find the major dimensions of linguistic variation of Twitter trolling, I used the short text version of MDA on the 4,182 trolling tweets measured for the presence or absence of 69 linguistic features that occurred in more than 5% of the tweets (see Appendix 1). The MCA returned 69 dimensions ($L \leq 138\ categories - 69\ linguistic\ features = 69$), and for each dimension each tweet and each category of a linguistic feature (presence or absence) was assigned a positive or negative coordinate and a value indicating its contribution to the dimension.

Prior to interpretation, the dimensions of Twitter trolling were compared to the dimensions of general English Twitter (see section 4.7.2). Given that both analyses have different feature sets, it was not possible to correlate the dimensions according to the categories of linguistic features; however supplementary tweets enables tweets to be projected onto dimensions of linguistic variation by assigning them a coordinate which positions them in accordance to their association to the sets of linguistic co-occurrence patterns. Through supplementary tweets, it is possible to compare the coordinates of the trolling tweets aligned with both sets of dimensions.

As a result, each trolling tweet was also measured for the presence or absence of the 63 linguistic features used in the MDA of general English Twitter in Chapter 5, and then the trolling tweets were specified as supplementary in the MCA of general English Twitter. These supplementary trolling tweets were assigned a positive or negative coordinate for each dimension of linguistic variation of general English Twitter, which located their position on the cloud of active general English tweets, and indicates how associated they are to the dimensions of linguistic variation across General English Twitter.

This means that each trolling tweet has been assigned a coordinate for each dimension of linguistic variation across both studies. For example, Table 28 reveals the coordinates for Example C (a tweet from the trolling corpus) with respect to the dimensions of linguistic variation for Twitter trolling for which it is an active tweet, as well as the dimensions of linguistic variation for general Twitter for which it is a supplementary tweet.

> **Example C**: @PetrePowder @Celebritygrbage @amandacarpenter @benshapiro Lol □ "collaborate" is freaking funny in this context. You meant "corroborate". You're uneducated, right? Thanks for the laugh, sparky.

Table 28: The coordinates of trolling tweet Example A for the Dimensions of Twitter Trolling and the Dimension of General Twitter

Example C	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
Dimensions of Twitter Trolling	0.088	-0.125	-0.264	-0.144	0.216
Dimensions of General Twitter	0.328	-0.148	-0.307	-0.211	0.222

Because each trolling tweet has been assigned a coordinate revealing how associated it is to either set of dimensions, the degree of similarity between the dimensions of linguistic variation across the two studies can be assessed according to the coordinates of the tweets. Specifically, the coordinates of trolling tweets for the dimensions of linguistic variation of general English Twitter were correlated to the coordinates of trolling tweets for the dimensions of linguistic variation of Twitter trolling. If the dimensions of linguistic variation are the same across general English Twitter and Twitter trolling then a strong correlation would be expected (r > .85).

Table 29: Correlation Matrix of each trolling tweet's dimension coordinates for the dimensions of general English Twitter (_TTGT) and for the dimensions of Twitter trolling (_TT)

	Dim.1_ TTGT	Dim.2_ TTGT	Dim.3_ TTGT	Dim.4_ TTGT	Dim.5_ TTGT
Dim.1_TT	0.99	0.06	0.03	-0.07	0.08
Dim.2_TT	-0.06	0.93	-0.05	0.05	-0.31
Dim.3_TT	0.01	0.00	0.94	-0.06	0.14
Dim.4_TT	0.04	-0.17	-0.12	0.8	-0.17
Dim.5_TT	0.02	0.1	-0.11	-0.01	0.65

Table 29 is a correlation matrix of each trolling tweet’s coordinates for the dimensions of linguistic variation of General English Twitter (Dim.*n*_TTGT) and each trolling tweet’s coordinates for the dimensions of linguistic variation of Twitter trolling (Dim.*n*_TT). This table shows that Dimension 1, 2 and 3 are very strongly correlated. So correlated in fact that these dimensions should more or less be the same when it comes to interpretation. Table 29 also shows that Dimension 2 of Twitter trolling is moderately negatively correlated to Dimension 5 of general Twitter. This will be addressed in section 6.6. In comparison to Dimension 1, 2, and 4, Dimension 4 is less correlated, although the correlation is still strong (r =.8), suggesting that the range of linguistic variation may differ slightly, potentially calling for a new interpretative label that represents the opposition more clearly. Finally, Table 29 shows that Dimension 5 is moderately correlated (r=.65), suggesting a new interpretative label.

Overall, the results of this preliminary comparison of the dimensions through the coordinates of the trolling tweets has indicated that at least the first three dimensions of general English Twitter are also major patterns of linguistic variation for Twitter trolling. Given that Twitter trolling is situated in the context of general Twitter, this is perhaps not completely surprising. Nevertheless, it is interesting to see that these patterns are so strong, especially considering that both corpora were collected with different constraints.

Table 30: Variances of Dimensions (eigenvalues and modified rates)

Dimension	1	2	3	4	5
Eigenvalue	0.112	0.032	0.025	0.023	0.02
Modified Rates	0.94	0.03	0.011	0.007	0.002

Following this comparison, I extracted the first five dimensions, as these were readily interpretable, and based on the modified rates of the eigenvalues, these first five dimensions explain a large portion of the variance (see Table 30). For each dimension, I interpreted the features that contributed above the average contribution $\left(\frac{100}{69\ linguistic\ features\ x\ 2\ categories} = \frac{100}{138} = \ 0.72\right)$ and the strongly contributing tweets. Those features and tweets contributing the most with positive coordinates were interpreted in opposition to the features and tweets with negative coordinates for the underlying communicative function. Each dimension is presented below in a separate section. More examples for each dimension can be found in Appendix 4.

6.1. Dimension 1: Length

The linguistic features most strongly contributing to Dimension 1 are presented in Table 31. Table 31 shows that positive Dimension 1 is characterised by the *presence* of 40 linguistic features, whereas negative Dimension 1 has the *absence* of 19 features, suggesting that Dimension 1 is opposing long tweets with short tweets, like Dimension 1 of General English tweets (see Chapter 5). In comparison to general English tweets, only URLs are present in the non-strongly contributing features on the negative side of Dimension 1. The reason for this is because the feature *emojis* was not included in this analysis, as its occurrence in the trolling tweets was below 5%.

Table 31: The linguistic features strongly contributing to positive and negative Dimension 1 of Twitter trolling (coordinates; contributions)

1	+	**Present features**: Other_Verb (0.354; 1.16), Attributive_Adjective (0.383; 1.12), Full_Stop (0.426; 1.408), Preposition (0.429; 1.533), Stative_Form (0.509; 1.687), Third_Person_Singular_Verb (0.549; 1.686), Other_Adverb (0.564; 1.76), Definite_Article (0.574; 1.753), Indefinite_Article (0.582; 1.514), WH-Word (0.583; 0.854), First_Person_Pronoun (0.59; 1.65), Past_Tense_Verb (0.607; 1.775), Possession (0.623; 1.594), Comma (0.633; 1.756), Contracted_Forms (0.637; 1.264), Subject_Pronoun (0.637; 2.379), Analytic_Negation (0.669; 1.857), Predicative_Adjective (0.672; 1.335), Public_Verb (0.676; 1.336), Coordinating_Conjunct (0.677; 2.059), Demonstrative_Pronoun (0.68; 0.908), Stance_Verb (0.693; 0.806), Progressive (0.701; 0.807), Nominalisation (0.702; 1.293), Private_Verb (0.718; 1.46), Third_Person_Pronoun (0.721; 1.911), IT (0.745; 1.398), Other_Subordinator (0.748; 1.095), Object_Pronoun (0.756; 0.942), Modal_Prediction (0.762; 1.05), Indefinite_Pronoun (0.771; 0.972), Infinitive (0.79; 1.844), Auxiliary_DO (0.803; 1.21), Relatives (0.815; 0.904), Contrastive_Conjunct (0.828; 1.049), HAVE_Main_Verb (0.86; 0.937), Passive (0.905; 1.009), Perfect_Aspect (0.916; 0.947), Complementation (0.952; 1.191), Conditional_Subordinator (1.05; 1.069).
	-	**Absent features**: Other_Noun (-1.229; 2.861), Other_Verb (-0.894; 2.93), Preposition (-0.783; 2.801), Full_Stop (-0.639; 2.11), Attributive_Adjective (-0.553; 1.617), Subject_Pronoun (-0.53; 1.979), Stative_Form (-0.517; 1.712), Other_Adverb (-0.423; 1.318), Third_Person_Singular_Verb (-0.419; 1.287), Definite_Article (-0.403; 1.231), Past_Tense_Verb (-0.362; 1.058), Coordinating_Conjunct (-0.361; 1.097), First_Person_Pronoun (-0.343; 0.96), Proper_Noun (-0.34; 0.724), Comma (-0.325; 0.902), Analytic_Negation (-0.317; 0.882), Indefinite_Article (-0.308; 0.8), Possession (-0.291; 0.744), Third_Person_Pronoun (-0.287; 0.759).

Table 32 presents the trolling tweets most strongly associated with positive Dimension 1 and Table 33 presents the trolling tweets most strongly associated with negative Dimension 1. Examples 1-5 in Table 32 are considerably longer than Examples 1-6 in Table 33, which are predominantly URLs.

Table 32: The trolling tweets most strongly associated with positive Dimension 1

	Tweet	Coord	Contrib
1	@BroadStBulliz @realDonaldTrump Oh yeah I'm sure he was lying despite not knowing he was filmed and potentially risk losing his job. It's been a year, there's no evidence and if there was he would've been in jail a long time ago. Once the DOJ acts, it's all downhill and screaming to the sky for you bud	0.87	0.16
2	@OneAboveTwo @seanspicer No he's not! If you've ever listened to @Sean Spicer talk about those in education who dedicate their lives to our children's future you'd hear he knows teachers don't get paid near what their worth to us!! I believe he's saying now that a settlement is reached, back to students.	0.86	0.16
3	@Taipan30 @BillPeriman @jpcronk @TomSeward5 @BluSthil @RyanAFournier This may come as a shock you all but just because news organizations are critical of your failing POTUS doesnt make them fake. Even if your POTUS tweets and says it so. Does that make sense? I can only hope knowing you all have been robbed of your ability for critical thinking	0.83	0.15
4	@FBI The higher ups in the FBI has ruined your reputation...they are criminals and need to be locked up..FYI, why hasn't Peter and Page been fired yet?? By the texts it's obvious that anyone else would have lost their jobs by now..	0.8	0.14
5	@margculbster @davidhogg111 I am now convinced you over estimate your own intelligence. How can one feel big and strong to belittle a kid if the person you are accusing of it, may be the same age as said kid? Thank you for the compliment. I disagree that I am just like Trump but hopefully when I grow up □ https://pbs.twimg.com/media/DXFVklJVQAEp-tn.jpg	0.8	0.14

Table 33: The trolling tweets most strongly associated with negative Dimension 1

	Tweet	Coord	Contrib
6	https://t.co/8AotPTWwZX	-0.634	0.086
7	https://t.co/HcA6pFbAT3	-0.634	0.086
8	https://t.co/jALOKkLk3x	-0.634	0.086
9	https://t.co/1ycUETCNzL	-0.634	0.086
10	□□ https://t.co/i2NAjVNgB5	-0.634	0.086

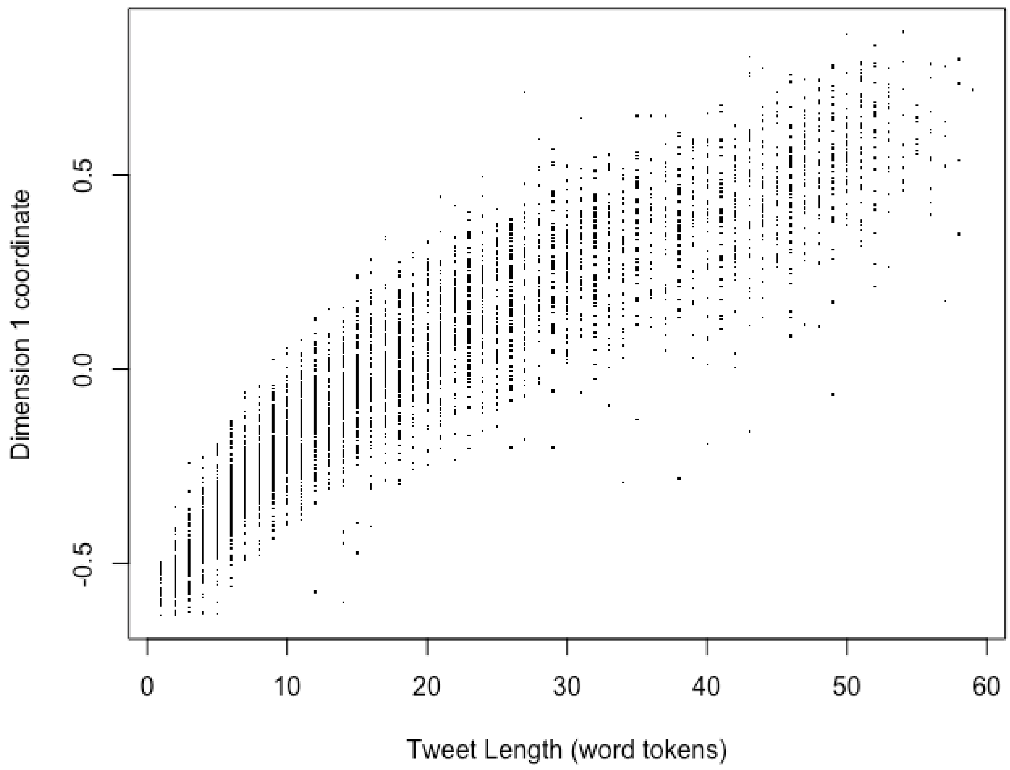
Like the analysis of general tweets, tweet length in word tokens was included as a supplementary variable to observe if tweet length was correlated to the dimension coordinates of tweets. Table 34 shows that tweet length is strongly positively correlated with Dimension 1. Figure 5 also shows this strong correlation, which demonstrates that as tweets get longer, their coordinate on Dimension 1 rises. The correlation of Dimension 1 to tweet length across trolling tweets is stronger than general English tweets, indicating that tweet length in this corpus of trolling tweets is the strongest influence on the presence of features. This may be due to the inclusion of more features in this analysis, enabling longer tweets to include the presence of more features, as well as trolling tweets on average being longer than general tweets.

Table 34: The tweets' dimension coordinates correlated to tweet length

	Dim.1	**Dim.2**	**Dim.3**	**Dim.4**	**Dim.5**
Tweet Length (word tokens)	0.912	0.148	0.016	-0.036	0.017

Overall, these results are equivalent to Dimension 1 of general English Twitter, and therefore like that analysis, Dimension 1 is excluded from further interpretation.

Figure 5: The trolling tweets Dimension 1 coordinates to length of tweet in word tokens



6.2. Dimension 2: Informational versus Interactive

The linguistic features most strongly contributing to Dimension 2 of Twitter trolling are presented in Table 35. Like General Twitter in Chapter 5, an opposition between an informationally dense style and an interactive style connects the linguistic features most strongly contributing to Dimension 2. Specifically, positive Dimension 2 is characterised by numerous noun types (*non-initial mentioning, proper noun, numeral noun*) and noun modifiers (*definite article, graduation, numeral determiner*), associated with the integration of information, as well as hashtags, associated with broadcasting the content to an audience beyond one's followers.

Table 35: The linguistic features most strongly contributing to Dimension 2 of Twitter trolling (coordinates; contributions)

2	+	**Present features**: Preposition (0.206; 1.222), Definite_Article (0.277; 1.413), Exclamation_Mark (0.396; 0.921), Proper_Noun (0.42; 4.062), Graduation (0.52; 1.002), URL (0.651; 3.774), Capitalisation (0.668; 4.833), Numeral_Noun (0.793; 2.673), Numeral_Determiner (0.798; 2.432), Hashtag (1.17; 6.001), Non-Initial_Mention (1.234; 3.581).
		Absent features: First_Person_Pronoun (0.173; 0.844), Contracted_Forms (0.181; 1.107), Second_Person_Pronoun (0.362; 3.263), Initial_Mention (1.104; 10.31)
	-	**Present features**: Contracted_Forms (-0.568; 3.481), Auxiliary_DO (-0.495; 1.591), Second_Person_Pronoun (-0.457; 4.121), Private_Verb (-0.406; 1.615), Question_Mark (-0.366; 1.225), Interjection (-0.35; 1.15), IT (-0.329; 0.945), Predicative_Adjective (-0.323; 1.068), Analytic_Negation (-0.314; 1.418), First_Person_Pronoun (-0.297; 1.452), Initial_Mention (-0.258; 2.405), Subject_Pronoun (-0.195; 0.77).
		Absent features: Proper_Noun (-0.446; 4.318), Preposition (-0.376; 2.233), Other_Noun (-0.372; 0.907), Attributive_Adjective (-0.23; 0.97), Capitalisation (-0.213; 1.543), Definite_Article (-0.194; 0.992), URL (-0.162; 0.937).

Table 36 presents the tweets most strongly associated with positive Dimension 2. All of the tweets in Table 36 are informationally dense and integrate specific details on particular referents in the short space of a tweet. For instance, in Example 11, the first sentence (26 words) integrates a huge amount of content. It describes that Donald Trump has played golf 1 of every four days that he has been president. It also details that there was a segment on @allinwithchris about Trump's activity. And finally, it explains that a pin on "@11thHour #MSNBC" was inspired by that particular segment.

Table 36: The trolling tweets most strongly associated with positive Dimension 2

	Tweet	Coord	Contrib
11	Tonight's pin on @11thHour #MSNBC was inspired by a segment on @allinwithchris tonight about Trump playing golf 1 of every 4 days he's been Pres. WH trying to hide this. If the Mueller investigation is allowed to continue, maybe it'll give 45 more time for his favorite activity. https://pbs.twimg.com/media/DSGXnSwV4AE2OUg.jpg	0.683	0.345
12	One of David Hogg 's personal friends stabs her newborn to death then dumps the body in a neighbor's shed and goes to sleep. #GunReformNow #GunControlNow #GunControlNever #nra #ObamaTookMillionsFromTheNRAToo https://t.co/TAJhlo5viw	0.658	0.32
13	These so-called "Journalist" need their credentials pulled immediately, and charged if any crimes were committed. #SorosPuppets #QAnon THE WIKILEAKS LIST: At Least 65 MSM Reporters Were Meeting with and/or Coordinating Offline with Top Hillary Advisors https://t.co/U802GhOqz9	0.655	0.317
14	#NewProfilePic Was @willowhalegreen but banned for the truth. 5.500 followers to zero because TWATTER loves Islam and hates patriots.we will never bow to the cult of death and destruction.. https://pbs.twimg.com/media/DRW_G_iW0AAfUJr.jpg	0.607	0.272
15	ANDREW BOLT: "Well, this is embarrassing. The crippled Turnbull Government is fighting for its life but has now benched its best two election campaigners: @TonyAbbottMHR and @Barnaby_Joyce." #BringBackAbbott #auspol https://t.co/jqNEdRJx6G	0.603	0.268

Additionally, the noun phrase "*One of David Hogg 's personal friends"* in Example 12 includes numeral noun "One", the preposition "of", the proper noun in possessive form "David Hogg 's", as well as the attributive adjective "personal" and the general noun "friends" to incorporate post-nominal modification of the numeral noun to signal the relation of belonging to a group (i.e. it is one of many). Importantly, the whole phrase could be replaced with a pronoun 'he/she/they' or even the name of the individual (Erica Gomez). However, pronouns tend to be used when there is a shared communicative context, as it is clear to whom the pronoun refers. The provision of detailed and specific descriptions, as opposed to pronouns, is therefore indicative of a lack of shared context between the reader and the author. Moreover, it is suggestive of communication that is produced when there is time to plan and edit (Biber, 1988), and in this case, identify particular relations.

By contrast, negative Dimension 2 is characterised by several features associated with a shared communicative context (*subject pronouns, first and second person pronouns*, *pronoun IT, interjections*), orality (*contractions*), a fragmented production of text, and the expression of stance (*predicative adjectives, private verbs),* characteristic of an interactive style (Biber, 1988).

Table 37: The trolling tweets most strongly associated with negative Dimension 2

	Tweet	**Coord**	**Contrib**
16	@Salon I'm white. I don't see it. Morons. How is he a racist? I don't think you know what racism is.	-0.527	0.205
17	@RomeDoesIt i don't know why you're still so concerned about a troll tweet. i hate idiots like you.	-0.493	0.179
18	@NEERAJ_AGARWAL_ @barmanamar1976 @CFBKEW @DeanKo @WeAreWakinUp @phiroc @_Gravity_Man @Theflateartherz @VerumBellator1 @FlatEarthCity @ItsFlatFolks @facebones777 @HomeoReikiDogs @SpeakToMeInDots @Its_Stationary @GodofGreen2 @nutsyLFC @Spacehehehe @ADalassio @VickyAlam18 @Th3NewMoon @BadBuc99 @ericdubay @IronRealmMedia @catomilla @IllCity_Luck @jeranism @TheWrongQuest @jaredvc @mode23 @hugh_bothwell Just because you don't understand something doesn't mean it isn't true lol.	-0.479	0.169
19	@idkasuri @Fabulous_IK @ummesalaar @Paracha_Pk @kambohguyforpti @StaunchInsafian @FauziaKasuri @BBhuttoZardari Oh we know what you're talking about. Get it off your chest anyway if you want.	-0.472	0.165
20	@helenlooise Did you have some point to make or whatever? I don't know what you're on about.	-0.462	0.158

Table 37 presents the tweets most strongly associated with negative Dimension 2, all of which are interacting with particular Twitter users through *initial mentioning* and *second person pronouns,* and encoding personal stance. Example 16, for instance, reacts to a previous tweet by signalling that they disagree through *analytic negation* and *auxiliary DO* (*I* ***don't*** *see it*). Example 16 also interacts with the addressee by asking them a question (***How*** *is he a racist****?***).

Overall, these results are equivalent to Dimension 2 of general English Twitter, which indicates that trolling tweets can be informational and broadcast to a particular audience, and that they can also be interactive.

6.3. Dimension 3: Personal versus Other Description

The linguistic features most strongly contributing to Dimension 3 of Twitter trolling are presented in Table 38. Corresponding to Dimension 3 of General English Twitter, an opposition between a personal style and a topic-focused style connects the linguistic features and tweets most strongly contributing to Dimension 3. In particular, positive Dimension 3 is characterised by *object pronouns*, especially *first* and *third person pronouns*, indicating that action has acted upon the author or other patients.

Positive Dimension 3 is also characterised by a variety of verb types and constructions (*private verbs, public verbs, stance verbs, phrasal verbs, infinitives, pro-verb DO, auxiliary DO, modal of possibility* and *passive* *constructions*), and *imperatives*, which, like general tweets, occur in the trolling tweets most associated to Dimension 3 to encode personal experience and stance, and make personal requests. Table 39 presents the trolling tweets most associated with positive Dimension 3. The author of Example 22 in Table 39 reports on their experience through the *object pronoun* (*Thomas Wictor has* ***me*** *blocked*). Additionally, Example 22 encodes the author's personal desire by making a personal request (*Please retweet Thomas or me*), and through the first person pronoun and stance verb "want" (***I want*** *this information exposed!*).

Table 38: The linguistic features most strongly contributing to positive and negative Dimension 3 (coordinates; contributions)

3	+	**Present features**: Private_Verb (0.269; 0.919), Analytic_Negation (0.271; 1.374), First_Person_Pronoun (0.277; 1.632), Infinitive (0.278; 1.025), Third_Person_Pronoun (0.299; 1.472), Public_Verb (0.348; 1.59), Phrasal_Verb (0.359; 1.028), Passive (0.37; 0.756), Modal_Possibility (0.39; 0.986), Question_Mark (0.395; 1.853), Pro-Verb_DO (0.413; 0.863), Conditional_Subordinator (0.42; 0.77), WH-Word (0.42; 1.995), Stance_Verb (0.526; 2.085), Imperative (0.56; 2.492), Auxiliary_DO (0.78; 5.134), Object_Pronoun (0.828; 5.081).
		Absent features: Contracted_Forms (0.161; 1.145), Attributive_Adjective (0.196; 0.909), Third_Person_Singular_Verb (0.216; 1.537), Predicative_Adjective (0.253; 2.855), Stative_Form (0.564; 9.176).
	-	**Present features**: Graduation (-0.928; 4.137), Predicative_Adjective (-0.851; 9.619), Stative_Form (-0.556; 9.045), Contracted_Forms (-0.507; 3.599), Demonstrative_Pronoun (-0.341; 1.027), Third_Person_Singular_Verb (-0.283; 2.014).
		Absent features: Other_Verb (-0.312; 1.599), First_Person_Pronoun (-0.161; 0.949), Auxiliary_DO (-0.133; 0.873), Object_Pronoun (-0.121; 0.744).

Positive Dimension 3 is also characterised by *conditional subordinators,* which often occur in the tweets to state a condition or hypothesis that is personal, such as Example 23 (***if** I wanted to say something to you @billings_steve I would have brought you up*).

Table 39: The trolling tweets most strongly associated with positive Dimension 3

	Tweet	Coord	Contrib
21	@VictoriaBanvil2 @DavidBegnaud What are you a fascist ? You decide if someone can reply or not? Don't confuse me with a colonial subject... that you tell what to do or not.	0.588	0.332
22	Since Thomas Wictor has me blocked (don't know why) I am tweeting this valuable information he has provided. Handwriting from year book Roy Moore accuser provided, does not match and person got Old Hickory House wrong. Please retweet Thomas or me. I want this information exposed!	0.582	0.325
23	@billings_steve @VigorousRaDiCaL @lorenzabraham12 @GhosTNinjaFtW then do me a favor don't say i said something about you when I didn't. If I wanted to say something to you @billings_steve I would have brought you up	0.582	0.325
24	@AnnCoulter @SUPgrlCaroline @realDonaldTrump Trump, with all the idiotic choices on immigration has done what the whole world could not do, he has turned me against him. Let Mueller lock him up for lying to us on the wall, and amnesty! At this point I really don't care if he gets impeached or locked up!	0.571	0.313
25	@ESPNFC When did USA become a team to miss @the world cup? Don't Make me laugh! They can miss 10 World cups for all I care! Won't be missed!	0.547	0.287

Question marks and *WH-words* are also associated with positive Dimension 3 and they occur to form questions, especially rhetorical questions, which allow the author to respond and encode their personal opinion or demand, such as Example 21 (***What** are you a fascist **?** You decide if someone can reply or not**?** Don't confuse me with a colonial subject [...]*) and Example 25 (***When** did USA become a team to miss @the world cup**?** Don't Make me laugh!)*.

Table 40: The trolling tweets most strongly associated with negative Dimension 3

	Tweet	Coord	Contrib
26	@BobbyMartin044 @TBama23 @NFLResearch Its the whitest team in America. since 9/11 the Patriots have been Unstoppable. New ENGLAND, PATRIOT, red-white-blue, the HAMPTONS, POLITICS, it's the whitest area of the country, the coach is white, QB is white, RB is white, WRs r even white! There is a white man on the helmet!!	-0.547	0.286
27	@AlterSol @brianherman @pma19722 @RightlyNews @Fuctupmind Facts matter: Trump has best stock market performance in 20+ years. Lowest unemployment in 17 years, um, oh yea, that's Better than Obama ... EVER!	-0.542	0.281
28	@thebuddhacat1 @IanMCohen @azmachman @John_AKA_Becker @SassBaller @CNN @BarackObama @POTUS Stupid at its best coming from a "nasty Buddha cat".	-0.49	0.23
29	@THR Well... The Room is pretty horrible... But at least it's better that #LastJedi https://t.co/r01rwV7o8f	-0.483	0.224
30	@breton_anne @DestinyandBruce @SergeFauchet @EvOConnor15 @TimRunsHisMouth @DonaldJTrumpJr @POTUS @realDonaldTrump @FLOTUS Nope. Obama wears the biggest liar title well. But that's okay. Sit back and enjoy the ride.	-0.475	0.216

By contrast, negative dimension 3 is characterised by features associated with characterising and describing subjects that are external to the self (see Table 38). In particular, stative forms (e.g. BE as a main verb and copular verbs) are used in the tweets associated to negative Dimension 3, presented in Table 40, to characterise and describe a subject, such as Example 26 (*the coach **is** white).* Additionally, negative Dimension 3 is characterised by *predicative adjectives*, such as “horrible” in Example 29 (*The Room is pretty **horrible***), and *superlative* and *comparative* forms, such as “biggest” in Example 30 (Obama wears the **biggest** liar title well) and “better” in Example 29 (But at least it's **better** that #LastJedi) to characterise and rank the subject. Finally, t*hird person singular verb forms* and *demonstrative* *pronouns*, such as in Example 27 with “has” and “that” are used to mark that the subject is not the author (*Trump **has** best stock market performance in 20+ years… **that**’s better than Obama*).

Overall, these results are exceptionally similar to Dimension 3 of general English Twitter, which indicates that trolling tweets can be personal and self-report, and that they can also be more topic-focused and characterise external entities.

6.4. Dimension 4: Promotional versus Oppositional

The linguistic features that are strongly contributing to Dimension 4 of Twitter trolling are presented in Table 41. Table 42 presents the trolling tweets most associated with positive Dimension 4. Whilst many of the features strongly associated with positive Dimension 4 of Twitter trolling are the same as those associated with General Twitter’s positive Dimension 4, there are some differences, including the strong association of *third person pronouns, contrastive conjunctions*, and *complementation*, all of which are associated with negative Dimension 4 of general Twitter.

Conditional subordinators and *modals of prediction* are also strongly associated with positive Dimension 4 of Twitter trolling; however *conditional subordinators* was not a feature in the general Twitter analysis, as it occurred in fewer than 5% of the tweets, and *modals of prediction* were associated with positive Dimension 4 of general Twitter, but not strongly (contribution = 0.2, coordinate = 0.2).

Table 41: The linguistic features strongly contributing to positive and negative Dimension 4 of Twitter trolling

4	+	**Present features**: Third_Person_Pronoun (0.214; 0.808), IT (0.286; 0.988), Stance_Verb (0.303; 0.741), Contrastive_Conjunct (0.367; 0.99), Complementation (0.372; 0.875), Modal_Prediction (0.389; 1.312), Conditional_Subordinator (0.415; 0.802), Subject_Pronoun (0.443; 5.532), Hashtag (0.455; 1.258), Contracted_Forms (0.462; 3.192), Predicative_Adjective (0.467; 3.095), First_Person_Pronoun (0.499; 5.685), Object_Pronoun (0.51; 2.059), Perception_Verb (0.519; 1.021), URL (0.814; 8.172), Non-Initial_Mention (0.858; 2.404).
		Absent features: Question_Mark (0.161; 1.283), Preposition (0.192; 0.805), Second_Person_Pronoun (0.198; 1.359), Attributive_Adjective (0.214; 1.159), Other_Verb (0.235; 0.973), Other_Noun (0.66; 3.967), Initial_Mention (0.832; 8.116).
	-	**Present features**: Question_Mark (-0.626; 4.976), Imperative (-0.544; 2.511), Synthetic_Negation (-0.471; 0.798), WH-Word (-0.352; 1.498), Nominalisation (-0.348; 1.525), Second_Person_Pronoun (-0.25; 1.717), Initial_Mention (-0.194; 1.893), Attributive_Adjective (-0.148; 0.803).
		Absent features: Subject_Pronoun (-0.369; 4.602), First_Person_Pronoun (-0.29; 3.307), URL (-0.202; 2.03), Contracted_Forms (-0.147; 1.015), Predicative_Adjective (-0.139; 0.918).

In general, these inconsistent features are associated with idea elaboration (Chafe, 1982; 1985). For example, *complementation* is often used to expand an idea unit, such as Example 35 in Table 42 (*you're too westernized to even know* ***what I'm talking about***). *Contrastive conjunctions* are used to show a contrast or difference, and thus similarly expand an idea unit by introducing an alternative position. Additionally, *conditionals* and *modals of prediction* are used for talking about possible and future events (Finegan, 1982), and can be used to elaborate on a personal idea or opinion, such as Example 32 (*It **would** be great **if people could refute things they had disagreement on with facts***).

In addition to features associated with idea elaboration, positive Dimension 4, like Dimension 4 of general Twitter, is characterised by features associated with promoting, including CMC features such as *hashtags, non-initial mentioning*, and *URLs*, which are used to gain attention, publicly refer to and promote others, and share additional content that the author perceives to be important. For example, the *hashtag* “#BYU” in Example 31 is used to direct the content of the tweet to people interested in Brigham Young University basketball team and gain their attention. Additionally, *non-initial mentioning* and *URLs* are used in Example 32 to promote @Boxerworks and share his tweet with the author’s followers.

Table 42: The trolling tweets most strongly associated with positive Dimension 4

	Tweet	Coord	Contrib
31	We're going undefeated! #BYU https://t.co/XeyK6ExhRk	0.622	0.396
32	I don't know if I s/b proud or disappointed that 5 people blocked me on my thread today. It would be great if people could refute things they had disagreement on with facts. It seems to be a high bar on Twitter. I appreciate @Boxerworks grace in recognizing a Twitter amateur. □ https://twitter.com/Boxerworks/status/956998687464488962	0.551	0.311
33	I'm hoping 10 months from now trump will be #OutOfOffice #ImpeachTrump https://t.co/v5hEvZWvnY	0.522	0.279
34	@marieclaire I think it's mostly because females are so weak, so it feels as painful as a heart attack to them. For men, it would just be closer to stubbing a toe.	0.514	0.27
35	lol you're too westernized to even know what I'm talking about. https://t.co/gnlQeyl7Vh	0.497	0.253

Positive Dimension 4 is also characterised by various forms associated with encoding personal stance, such as *stance verbs* and *perception verbs*, *predicative adjectives* and *complementation,* as well as pronominal forms in *object* and *subject form*, especially *first* and *third person personal pronouns* and *pronoun IT,* suggesting an involved communicative style (Chafe and Danielwicz, 1987). These features are used in the tweets (Examples 31-35) to encode personal stance, especially on other people. For example, the author of Example 33 encodes their personal desire that they hope Trump will be impeached (***I'm hoping*** *10 months from now trump will be #OutOfOffice #ImpeachTrump*). Example 34 includes the author’s personal stance on women (***I*** *think it's mostly because females are so* ***weak***), and Example 35 encodes an opinion on the second person pronoun (*lol you're too* ***westernized*** *to even know* ***what I'm talking about***).

The tweets most associated with positive Dimension 4 in Table 42 are often encoding a personal viewpoint, and describing a possible or future event. However, unlike positive Dimension 3, these tweets are principally focused on gaining attention and showing off, which is achieved by broadcasting the content to particular feeds through *hashtags*, and sharing particular content through *URLs*. In particular, many of the tweets associated with positive Dimension 4 are quoting the tweets of other people and even their own. When the authors quote other people’s tweets, as opposed to responding to them directly, the author chooses to broadcast and show off their response to their followers and beyond. Moreover, when they quote their own tweets, they are aimed at getting other people to notice the tweet.

Table 43: The trolling tweets most strongly associated negative Dimension 4.

	Tweet	Coord	Contrib
36	@BradfordNims @PaulsEgo Why can't you and all the pro gunners just acknowledge the reality of your position? You don't care about people being gunned down in churches, schools, and other public places. You think your personal "Freedom" to have an AR15 trumps the rights of the others to live. Be a man.	-0.537	0.295
37	@DonSather2 @AustenLied @WelterPeggy @andrewcockerpoo @peterdaou Hey bro, what voters were purged? Hillary's? Burnie's? Trumps? Yours? Sit down little man.	-0.449	0.206
38	@SenWarren @LCVoters so why did you vote for the bloated, unaudited military budget - 1/4 of which could have funded repair of the all the problems in this country - homelessness, student debt, free college tuition, safe food, safe water, safe air to breath. You bloviate all the time. End the wars?	-0.428	0.187
39	@BorchidJoseph @BrentBozell ...And your EVIDENCE IS??? The "word" of ppl who WORK FOR demorats-and who WAITED 38 YEARS TO UTTER A WORD?? LOL. NOT evidence--VERY LATE HEARSAY BS!!	-0.415	0.177
40	@VictoriaBanvil2 @DavidBegnaud Right here and now...? Is PROMESA signed by Obama, which gave Puerto Rico an undemocratic US colonial junta. You think you can forget the murders, experiments, abuses, exploitation of lands for 119 years. You can't whitewash the shameful history of US colonialism...	-0.414	0.175

Alternatively, the features most strongly associated with negative Dimension 4 in Table 41 are used in the tweets presented in Table 43, in order to oppose and challenge a particular stance or entity. Apart from *imperatives* and *synthetic negation*, all features strongly contributing to negative Dimension 4 of Twitter trolling are also strongly contributing to negative Dimension 4 of general Twitter, and are similarly used in the trolling tweets to challenge and oppose a particular Twitter user. For example, question features (*WH-words, question marks*), *initial mentioning* and *second person pronouns* are used in Example 36, 37 and 39 to contest a particular position of another Twitter user and target them specifically.

Imperatives are used to demand action, *and synthetic negation* is used to negate something. Both forms are used in the tweets to express disagreement. For example, the imperative in Example 37 simultaneously expresses disagreement in the addressee's position and silences them (*Sit down little man*). *Synthetic negation* negates a particular action, but it does so in a more emphatic way than *analytic negation*. For instance, compare the following two tweets expressing the same content. Example D has *analytic negation*, whereas Example E has *synthetic negation*:

D. *@JackSun01 @Trumpism_45 @realDonaldTrump @POTUS So if you were illegally spied on and persecuted for over 2 years when there* ***isn't evidence*** *against you*, *while the MSM smears your good name on a daily basis, you'd just sit quiet and never speak up for yourself?[...]*

E. *@JackSun01 @Trumpism_45 @realDonaldTrump @POTUS So if you were illegally spied on and persecuted for over 2 years with* ***no evidence*** *against you, while the MSM smears your good name on a daily basis, you'd just sit quiet and never speak up for yourself?[...]* (see Appendix 4; Dimension 4 coordinate: -0.36, contribution: 0.134).

Example E with synthetic negation states that there is not any evidence about Trump much more categorically and emphatically, suggesting that the tweets with synthetic negation often oppose and challenge statements vigorously.

Overall, the linguistic features and the tweets most strongly associated with negative Dimension 4 are connected by an underlying communicative function to oppose and challenge. Despite the discrepancies between the features associated with Dimension 4 of Twitter trolling and general Twitter, an examination of the linguistic features and their use in the trolling tweets indicates that they have the same underlying communicative function. Consequently, Dimension 4 of Twitter trolling is interpreted as representing an opposition between trolling tweets used to promote and gain the attention of others with tweets that are used to oppose and challenge other people.

6.5. Dimension 5: Incivility versus Civility

The linguistic features that are strongly contributing to Dimension 5 are presented in Table 44. Table 45 presents the tweets most strongly associated with positive Dimension 5. Positive Dimension 5 is characterised by interactive features, such as WH-words, question marks, and second person pronouns, which are used in the tweets to ask adversarial and abusive questions, such as Example 41 (***Do you*** *know* ***what*** *a question mark is, stupid****?****).*

There are also features associated with characterising and describing a subject, such as *stative forms, predicative adjectives, demonstrative pronouns,* and *demonstrative determiners*. These features often occur in the tweets to accuse and insult, such as Example 41 (You **are** a brainless trump cult worshipper) and Example 43 (***This*** *is your problem*).

Table 44: The linguistic features most strongly contributing to positive and negative Dimension 5 of Twitter trolling

5	+	**Present features**: Stative_Form (0.153; 0.845), Third_Person_Singular_Verb (0.182; 1.024), Predicative_Adjective (0.23; 0.864), Second_Person_Pronoun (0.24; 1.829), Capitalisation (0.251; 1.092), Imperative (0.285; 0.798), Demonstrative_Pronoun (0.315; 1.084), Progressive (0.346; 1.093), URL (0.367; 1.919), Profanity (0.37; 0.83), Demonstrative_Determiner (0.377; 0.853), Private_Verb (0.479; 3.614), Hashtag (0.608; 2.6), Question_Mark (0.79; 9.18), WH-Word (0.949; 12.588), Non-Initial_Mention (1.066; 4.284), Complementation (1.075; 8.432).
		Absent features: Past_Tense_Verb (0.134; 0.811), Third_Person_Pronoun (0.151; 1.17), Full_Stop (0.167; 0.802), Subject_Pronoun (0.171; 1.146), Initial_Mention (0.47; 2.998).
	-	**Present features**: Passive (-0.662; 2.999), Perfect_Aspect (-0.612; 2.351), Gerund (-0.487; 1.209), Pro-Verb_DO (-0.438; 1.199), Contrastive_Conjunct (-0.415; 1.46), Place_Adverb (-0.411; 0.72), Third_Person_Pronoun (-0.38; 2.945), HAVE_Main_Verb (-0.342; 0.822), Past_Tense_Verb (-0.225; 1.36), Subject_Pronoun (-0.206; 1.378).
		Absent features: WH-Word (-0.23; 3.047), Question_Mark (-0.204; 2.366), Second_Person_Pronoun (-0.19; 1.448), Stative_Form (-0.155; 0.857), Third_Person_Singular_Verb (-0.139; 0.782), Private_Verb (-0.135; 1.015), Complementation (-0.122; 0.956).

Positive Dimension 5 is also characterised by features associated with an aggressive and uncivil style including *profanity* and *imperatives*, which are often used in the tweets to personally attack their addressee and make demands of them, especially one's to silence them, such as Example 43 (*Mind your* ***fucking*** *business*), Example 44 (get it through your thick skull because I'm only saying this once more this is a stan account) and Example 45 (*Go back to Facebook with the rest of your kind*).

In addition, *non-initial mentioning, hashtags,* and *URLs* are also strongly contributing to positive Dimension 5. These features are often used in the tweets to publicly shame their addressee, such as Example 42 (*Why isn't* ***@brianstelter****, whose beat is covering the Media, reporting on this?)*, Example 43 (*So what should we do* ***@AnaKasparian*** *start a war with the Philippines? [...]* ***#TYTLive***), and Example 41 in the quoted tweet via the *URL*.

By using these features, the message does not only reach the addressee, but it is also broadcast to the author's followers, and followers of the hashtag. Overall, the tweets strongly associated to positive Dimension 5 tend to display an uncivil and accusatory style.

Table 45: The trolling tweets most strongly associated with positive Dimension 5

	Tweet	Coord	Contrib
41	Do you know what a question mark is, stupid? How's this for a new tactic: You are a brainless trump cult worshipper and yo mama is a bitcoin dumpster ho. https://twitter.com/SpecialEDxx/status/934095239890976768 …	0.6	0.426
42	Why isn't @brianstelter, whose beat is covering the Media, reporting on this? Hate me all you want, that's fine, but is this OK, Brian?	0.552	0.36
43	So what should we do @AnaKasparian start a war with the Philippines? This is your problem. Mind your fucking business. #TYTLive	0.548	0.356
44	@USAF_Frye @kmassey04 @TomiLahren @foxandfriends That's why your a single trump supporter... and I'm not fake my followers know what I really look like get it through your thick skull because I'm only saying this once more this is a stan account.... pic.twitter.com/QMKqPwRxbQ	0.519	0.319
45	@LexiHunt00 @mfleming877 @cathy58444301 @adarondax1 @Edible3Ball @realDonaldTrump @caramastrey @flowers3712 @SandraJH13 @cindyknoxville @marklevinshow @wishgrantlotus @CalebEatsBacon @California4Trmp @POTUS @MissTigerAngel @Greggorj @Cyptocon70 @BrutalVeracity @jaydixson Report what?. You don't like it.??? Then leave Twitter. This is our place now. Go back to Facebook with the rest of your kind. Buh Bye!!! @BarackObama @POTUS44 @MichelleObama @ObamaFoundation https://pbs.twimg.com/media/DXG4bYaW0AAehCM.jpg	0.487	0.281

By contrast, the features strongly contributing to negative Dimension 5 are much more associated with a polite and civil argumentative style. Many of the features are associated with narrativity, such as *subject pronouns*, especially *third person pronoun*, *past tense*, *passive construction*, and *perfect aspect*, which are used in the tweets, presented in Table 46, to report on past actions, especially of other people in order to critique them, such as Example 46 (***he's been*** *well coached on gun control ideology*), Example 49 (we ***were not shown*** the same solidarity) and Example 50 (*we* ***missed*** *an opportunity to have an important and necessary conversation*).

Table 46: The trolling tweets most strongly associated with negative Dimension 5

	Tweet	**Coord**	**Contrib**
46	@Loc8YourDignity @cookiebaker57 @thehill and "serviced" v "survived" but I got your meaning. Of course he has emotional knowledge and clearly he's been well coached on gun control ideology, when he goes off script he can get into the weeds. He's a 17 yo kid, no matter his o/ward maturity, he's still developing cognitively.	-0.432	0.221
47	@SuperSteveDV @SthrnMomNGram @karinagw @HebrewNational You have no prof he did it and don't even want to give him a chance. If there is prof then he deserves to be punished but until such time he remains innocent	-0.424	0.213
48	@floweraldehyde On the contrary, Apart from my school education which totalled to around 2 lakh from Kindergarten to Intermediate, I didnt receive any money for UG (around 1 cr paid by taxpayers). My PG also will be paid by taxpayer (most probably) around 2 crore rupees.	-0.422	0.21
49	@shannonrwatts Wealthy white Bernie supporters are obnoxious. Like Sarandon, they have other places to live and the means to do so, unlike those of us stuck here under Trump. HRC voters would've voted Sanders, we were not shown the same solidarity.	-0.421	0.21
50	@that_nocoiner @indystar We have no moral high ground here and we missed an opportunity to have an important and necessary conversation. Instead, folks like you can feel vindicated for not liking "those people" because, well, they don't like us either. https://t.co/YK235qMAui	-0.412	0.201

Other features associated with negative Dimension 5 include *HAVE as a main verb*, *pro-verb DO* and *gerunds*, which are used in the tweets to elaborate on a previously mentioned event or statement in order to provide an alternative point or reason for disagreeing, such as Example 47 (*You **have** no prof he **did** it […] If there is prof then he deserves to be punished but until such time he remains innocent*), and Example 50 (folks like you can feel vindicated **for not liking** "those people" because, well, they don't like us either). Alternative information and positions are also introduced through contrastive conjunctions, such as “but” in Example 46 and 47, “On the contrary” in Example 48, “unlike” in Example 49, and “Instead” in Example 50, which is also strongly associated with negative Dimension 5. Importantly, in comparison to the tweets on the positive side of the dimension, the tweets strongly associated with negative Dimension 5 are far less antagonistic and abusive, tending to provide corrections, alternative points of view, and offer critique in a civil manner.

Overall, Dimension 5 finds a distinction between trolling tweets that are uncivil and overtly antagonistic and abusive with trolling tweets that are non-abusive but instead are characteristic of civil argumentation. This dimension of linguistic variation was moderately correlated to general English Twitter because many of the features are the same, especially the kinds of narrative features on the negative side and the interactive features on the positive side of both dimensions. However, there also are some clear differences, such as *profanity* that occurs on the positive side of Twitter trolling, but on the negative side of general Twitter. Importantly, in the context of the trolling tweets the interactive features co-occur with alternative features and are exploited for abusive means, and the narrative features co-occur with alternative features to report on past events in order to politely correct or critique a previous point.

6.6. Discussion

With *trolling* being used to label a wide-range of behaviours and styles, this chapter sought to investigate whether the diversity of behaviours captured by the term is actually reflected in linguistic distinctions across numerous instances of trolling. Specifically, this chapter aimed to provide a linguistic description of trolling by identify and describing the major patterns of linguistic variation in a corpus of trolling tweets by applying the short text version of MDA to a corpus of 4,182 trolling tweets.

Given that Twitter trolling is situated in the context of Twitter, it is possible that trolling tweets could just be drawing on the major linguistic repertoires of general tweets. As a result, the major patterns of linguistic variation of trolling were compared to those of tweets more generally, established in Chapter 5 by correlating the coordinates of trolling tweets according to the Dimensions of Twitter Trolling with the coordinates of trolling tweet projected onto the Dimensions of general English Twitter as supplementary.

This comparison revealed that the first three dimensions were exceptionally correlated. This is a major finding, revealing that at least the first three dimensions of linguistic variation of general tweets are shared across general tweets and trolling tweets. It was also shown that Dimension 4 was strongly correlated, although it was less strong, and Dimension 5 was only moderately correlated.

After this preliminary comparison, the 5 dimensions of linguistic variation were interpreted for their underlying communicative function (see Table 47). Like General English Twitter, the first dimension of Twitter trolling reflected text length. The correlation between tweet length and the tweets' Dimension 1 coordinates was far stronger across Twitter trolling than general English Twitter, potentially because trolling tweets are on average longer (24 words) than general English tweets (16 words), and also because the analysis included more features (63 versus 69 linguistic features), as more features occurred in more than 5% of the tweets.

Also like General English Twitter, the second dimension of Twitter trolling opposes trolling tweets that are informationally dense with trolling tweets that are interactive. This supports previous research that has found that trolls post non-directed information or content, although this tends to be exaggerated or false in order to deceive, mislead, manipulate opinions or provoke a response (Donath, 1999; Hogan, 2012; Mihaylov et al., 2015; Lewandowsky et al., 2017; Netlingo.com, 1995; Herring et al., 2002). Many of the tweets in Table 36 associated with positive Dimension 2 are broadcasting and spreading press reportage. Specifically, they often include the headline of the news article, a comment on the content by the author, and the URL where the original article can be accessed. Examples 12, 13, and 15 are commenting on and broadcasting a piece of press reportage in this way. Although it is not possible to tell the veracity of all these reports and comments, Example 12 stands out as false and misleading, which claims that

David Hogg and Erica Gomez, who has stabbed her newborn child, are personal friends (*One of David Hogg ‘s personal friends stabs her newborn to death then dumps the body in a neighbor's shed and goes to sleep*). There appears to be, however, no relation between David Hogg and Erica Gomez. Thus, whilst determining the veracity of the content of these tweets is beyond the interpretation of the linguistic co-occurrence patterns, the results demonstrate that trolls can draw on this common linguistic repertoire in order to spread fake news and misleading information.

Moreover, this dimension supports previous research that has observed how trolls can be interactive and target their communication to particular individuals (Morrissey and Yell, 2016). One aim of some kinds of trolling is to obtain a reaction, and this can be achieved by employing an interactive style and by asking questions, such as Example 20 (@*helenlooise Did you have some point to make or whatever?*), which demands a response from the addressee. Another way to obtain a reaction is to antagonise the addressee and mock them, as this may provoke a response. For example, the troll tweet in Example 19 in Table 37 employs an interactive style and mocks their addressee by implying that they are emotional and that they might want to get something off their chest, as well as instructing them to do so *(@idkasuri @Fabulous_IK @ummesalaar […] Oh we know what you're talking about. Get it off your chest anyway if you want*).

Table 47: Summary of the Dimensions of linguistic variation of Twitter trolling

Dim		Summary
1	Long tweets vs. Short tweets	• This dimension opposes the presence of features with the absence of all features. • Trolling tweets' Dimension 1 coordinates are strongly positively correlated to text length. • The length of tweets is the strongest influence on the presence of features – the more words a trolling tweet has the more likely it will have the presence of features.
2	Informational broadcast vs. Interactive	• This dimension opposes informational broadcasting tweets with interactive tweets. Informational tweets consist of many features associated with the careful integration and broadcast of information, such as various noun types, noun modifiers, URLs and hashtags. Interactive tweets consist of interpersonal and involved features used to interact with other Twitter users like initial mentioning, pronouns, stance-related features, and features associated with a shared communicative context, such as interjections. • This may support previous research that suggests that trolls often aim to disrupt the conversation and aim to provoke a response. They can disrupt a conversation by interacting with people in a Twitter conversational thread. Additionally, they can post misleading or exaggerated integrated and informational content, which may provoke people to correct or point out the misconception. Additionally, trolls can ask provocative questions, demanding a response. Thus, they are dialogically provocative. • The analysis finds that an informational style is employed in order to spread fake news.
3	Personal vs. Other Description	• This dimension opposes trolling tweets that are highly personal and report on the self about what one is doing thinking or feeling with trolling tweets that are characterising and describing entities external to the self. • This is at variance to previous research that suggests that trolls need to be detached and remain emotionally divested because they tend to exploit the emotional sensitivities of others and thus by being emotional or revealing personal information they open themselves up to being trolled (Phillips, 2011).
4	Promotional vs. Oppositional	• This dimension opposes trolling tweets that are promotional and often seek to gain visibility with tweets that are opposing and challenging another Twitter users particular statement. • Trolling tweets that are displaying the promotional style are often expressing support for particular entities, which is contrary to previous research that has suggested that trolls are inherently oppositional (Phillips, 2016). • The promotional communicative style supports previous research that has described trolls to be attention-seeking (Hardaker, 2010; Cruz et al., 2018).
5	Incivility vs. Civility	• This dimension opposes trolling tweets that are highly abusive and uncivil with trolling tweets that employ a polite argumentative style. • Whilst this dimension is partially supported by previous research that has found that trolls are exceptionally hostile and abusive (), the civil argumentative nature is a strategy that has not been described nor identified in previous studies of trolling. • The civil communicative style employed by trolls may be a strategy to disguise their trolling intentions and provoke their victims to respond.

Additionally, trolls also seek to disrupt the conversation (Ansong et al., 2013). One way to do this can be to insert oneself into a conversation thread and interact with the other Twitter users. For example, there are two or more @usernames in the initial position in Examples 18 and 19, which indicates that there has been some form of conversation between these other mentioned Twitter users before this particular tweet. Although accessing the full conversation thread is not always possible to determine if the troll had in fact inserted themselves into the conversation, we can assume, for the most part, that the troll may have joined the conversation by replying to a particular tweet, and because all of these tweets have been identified and accused as trolling by other Twitter users (see section 3.1), the trolls have arguably disrupted the flow of the conversation. Overall, this dimension pattern reveals that trolls draw on the linguistic repertoires for broadcasting information and interacting in order to troll.

At the same time as being strongly correlated to Dimension 2 of general English Twitter, Dimension 2 of Twitter trolling is also slightly negatively correlated to Dimension 5 of general English Twitter, which reflect tweets that have a persuasive style with tweets that are more narrative. Many of the linguistic features strongly contributing to the persuasive side of the dimension are associated with interactivity and are shared across the dimensions (e.g. *question marks, second person pronouns, interjections, private verbs* and *auxiliary DO*). An interactive style in the persuasive tweets may be effective as it can imply friendliness and intimacy (Wells, Moriarty, and Burnett, 2006; Cui and Zhao, 2013) and may increase trust, which may persuade the reader to do something. By contrast, these features are not used in the trolling tweets to reduce the distance between the addressees and imply friendliness. In fact, these tweets create distance and imply unfriendliness by mocking and insulting the addressees. Despite this, the trolling tweets are arguably aiming to bring about something and persuade the addressee to do something. Specifically, they are often dialogically persuasive – aiming to provoke a reaction and therefore persuade the reader to respond. The negative correlation of Dimension 2 of Twitter trolling to Dimension 5 of general Twitter can offer linguistic support to research suggesting that trolls aim to provoke a reaction.

The third Dimension of Twitter trolling was also exceptionally correlated to the third Dimension of General Twitter. This dimension was similarly interpreted as opposing trolling tweets that were highly personal and self-reporting with trolling tweets that were more topic-focused and described other entities. This dimension indicates that trolling tweets vary according to how personal or detached they are, which runs counter to previous research that has suggested that trolls remain detached and avoid revealing any personal attachments or feelings because trolls exploit the emotional sensitivities of others (Phillips, 2011). In theory, trolls should remain detached to protect themselves from being targeted by trolling (Phillips, 2011). The trolling tweets strongly associated to positive Dimension 3 are perhaps counterintuitive for trolls because employing a personal style may make the trolls a target for other trolls. Nevertheless, as the third major dimension of linguistic variation, it suggests that trolls are perhaps more concerned with trying to achieve something else with this communicative style than being worried about being the target of trolls.

Specifically, the unexpectedness of this communicative style may work in the trolls' favour, especially as a deceptive strategy. Donath (1999) describes how trolls will try to deceive the community into thinking they are a legitimate community member, and then they will attempt to disrupt the community, whilst trying to keep up the facade of appearing as a genuine member. Given the similarity of these personal trolling tweets to general English tweets, these particular tweets may be purposeful in trying to convince their target that they are genuine and not trolling in order to be more successful in provoking a response.

However, not all trolls aim to provoke a reaction. The personal communicative style is used for a variety of purposes including to make personal requests, like in Examples 21, 22 and 23. In particular, the requests in 21 (*Don't confuse me with a colonial subject… that you tell what to do or not*) and 23 (*do me a favour don't say I said something about you when I didn't*) are essentially telling their addressee to take back what they have said before and shut up, which is the complete opposite to provoking a reaction. Overall, this dimension shows that trolling tweets are not always focused on remaining detached and emotionally divested, and actually do self-report and reveal personal opinions for a variety of communicative goals.

Alternatively, the trolling tweets associated with a topic-focused style are often describing and characterising subjects very matter-of-factly and sensationally. For instance, Example 30 categorically states that Obama is the biggest liar out of other Presidents of the United States (*Nope. Obama wears the biggest liar title well.*). This is an exaggeration, as it is impossible to know which President is the biggest liar for definite. Stating opinions as if they are matters of fact without any hedges, testimonials or even justifications may provoke another user to point out the misconception or correct the individual’s exaggeration, more so than if it was stated with these argumentative devices. Thus, whilst a topic-focused communicative style is a way to remain detached, providing partial support to previous research (Phillips, 2011), this communicative style is sometimes employed provocatively, which may also supports the notion that some trolls post hyperbolic or false statements in order to provoke another person to point out the correction (NetLingo, 1995-2015).

In addition to being provocative, the trolling tweets that are characterising and describing other entities are often doing so in order to correct an addressee and make an argument. For example, Example 27 characterises the stock market since Trump has been president as a way to correct their addressee and point out how biased they are towards Obama *(… um, oh yea, that's Better than Obama ... EVER!*). Thus, this dimension shows that characterising and describing entities is an important linguistic repertoire for trolling on Twitter and that this can be employed for a variety of purposes.

Like general English Twitter, the fourth dimension of linguistic variation found across Twitter trolling contrasts tweets functioning to promote with tweets functioning to oppose. This dimension also runs counter to previous research that has suggested that trolls are inherently oppositional (Phillips, 2016). Whilst this dimension shows that trolls can be oppositional, they are not *always* oppositional. In fact, many of the trolling tweets associated to positive Dimension 4 are publicly supporting particular people, ideas or ways of doing something, like Example 31, which shows support for Brigham Young

University basketball team by implying that they are going to win all their upcoming games. Example 32 also expresses appreciation towards another Twitter user. Moreover, these emotional expressions of support also run counter to the idea that trolls need to be emotionally detached (Phillips, 2011). Given mainstream media reports that emphasise the aggressive and hostile kinds of trolling posts, this linguistic repertoire is at variance to this. Tweets like example 31, for example, are relatively innocuous and characteristic of the kinds of talk of any sports fan before a sporting event, which raises the question about *why* particular tweets would be regarded as trolling, when they are expressing support for things.

By looking at the tweets more closely, it was observed that many of the things being supported could be negatively marked by others and construed as offensive. Example 33, for instance, is provocative to Trump supporters, and Example 34 is provocative to women (*I think it's mostly because females are so weak, so it feels as painful as a heart attack to them*). Overall, these expressions of support are arguably subtler and covert forms of trolling than the oppositional trolling tweets on negative Dimension 4, which are more targeted and antagonistic.

The fact that trolling tweets are often promotional and draw on resources for gaining attention supports previous research that has suggested that trolls troll for attention-seeking purposes (Hardaker, 2010; Cruz et al., 2018). Cruz et al. (2018) found that trolls targeted popular community members to gain the attention of the community and elicit reactions from them. The trolling posts associated to positive Dimension 4 increase noticeability by similarly targeting other Twitter users, but in a public way through quoting their tweets and hashtagging. Additionally, they may grab the attention of particular people by posting provocative or alternative things, like in the expressions of support for people, entities and ideologies. Overall, Dimension 4 could be conceptualised as a continuum of subtle and covert trolling posts to more explicit and overt trolling posts.

Finally, the fifth dimension of Twitter trolling opposes trolling tweets that are uncivil and abusive with trolling tweets that exhibit a civil argumentative style. This dimension of linguistic variation finds that trolling is not always abusive and hostile, but that it can actually provide fair critique and a reasoned and civil argument, which runs counter to the media's depiction of trolling. Whilst previous research has acknowledged that trolling is oppositional (Phillips, 2016), the degree of civility within that opposition has not been considered with it largely being characterised as uncivil. Although these posts have been perceived as trolling, a more polite and civil line of argumentation may be less expected as a trolling linguistic repertoire. Given the major negative depictions of trolls in the media, this could also be an interesting trolling strategy for deceiving their interlocutor into replying. Because the advice is to not engage with trolls, as they crave a reaction, trolls must deceive the community into thinking that their message is genuine so that they obtain a reaction (Donath, 1999). It could therefore be argued that a civil, non-abusive communicative style may be perceived as more genuine and less trollish than an overtly abusive and antagonistic one, especially as such a communicative style has not been described as a trolling linguistic repertoire before.

At the same time, not all trolls are trying to deceive their interlocutor into replying. A civil communicative style may be more effective for making a point more persuasive as their point is presented calmly and fairly. This can be useful in trying to spread misinformation or discredit a person, such as Example 46, which negatively presents David Hogg as someone who is a bit stupid (*he’s still developing cognitively)* and can only speak informatively when it is scripted (*he's been well coached on gun control ideology, when he goes off script he can get into the weeds)*. On the whole, this dimension reveals that civil and uncivil argumentation styles are important communicative styles for trolling on Twitter.

6.7. Conclusion

Overall, the analysis reveals many linguistic repertoires of Twitter trolling, which are at variance to previous research. Specifically, this chapter finds that trolling tweets are not always describing other things and are not always emotionally detached and divested, but can in fact display a very personal, involved and self-oriented communicative style. Additionally, this chapter finds that trolling tweets are not always oppositional but can actually be expressing support for particular things. Finally, this chapter identifies that in addition to uncivil communicative styles, trolling tweets can also be very civil and offer reasoned arguments. Whilst these patterns of linguistic variation have been surprising given previous research and the media’s representation of trolling, they could also simultaneously be obvious linguistic realisations of the troll’s major aim to deceive their victims into thinking that their tweeting intentions are genuine in order to provoke a reaction.

Cruz et al. (2018), for example, theorise that trolls firstly learn the behaviours of the community in order to assimilate towards them and appear genuine. Having learned these behaviours and been accepted into the community, they subsequently transgress and seek to disrupt the peace. Given the similarity between the major linguistic repertoires of general English Twitter and Twitter trolling, the findings from this chapter could offer support to this theory that trolls are drawing on similar repertoires as ordinary tweets in a process of assimilation. Specifically, they may be posting tweets that are highly interactive, informational, personal, impersonal, supportive, promotional and oppositional in order to disguise their trolling intentions and align with ordinary Twitter users so as to appear genuine in order to provoke a response from their victims. Additionally, trolls may be employing unexpected communicative styles, like a civil argumentative style in order to appear genuine.

Nevertheless, not all trolling aims to provoke a response. In fact, the analysis finds that some trolls are trying to silence their addressees and make them shut up. Additionally, the analysis finds that some linguistic repertoires are employed to spread misinformation and make it appear more genuine and persuasive. Moreover, the analysis finds that some linguistic repertoires are used to make a point, as well as to criticise, mock, taunt and insult their addressee. Thus, whilst some linguistic repertoires may be employed as a strategy for deceiving their addressees, this is not the primary aim for all trolls in each trolling tweet. Rather the results show that these linguistic repertoires are important for trolling on Twitter in order to achieve a variety of communicative goals.

Whilst this chapter shows that trolling tweets and general English tweets have a distinct linguistic repertoire, the analysis finds that trolling tweets and general English tweets draw on many of the same linguistic repertoires, although largely for divergent ends. This raises the next question addressed in the following chapter about how trolling tweets compare to general English tweets. Specifically, whilst we know that trolling tweets are a kind of tweet and that they share some linguistic repertoires, we know very little about which tweets are most similar to trolling tweets and what kinds of linguistic repertoires of general tweets do trolling tweets align with the most.

7. Comparing Trolling tweets along the Dimensions of general English Twitter

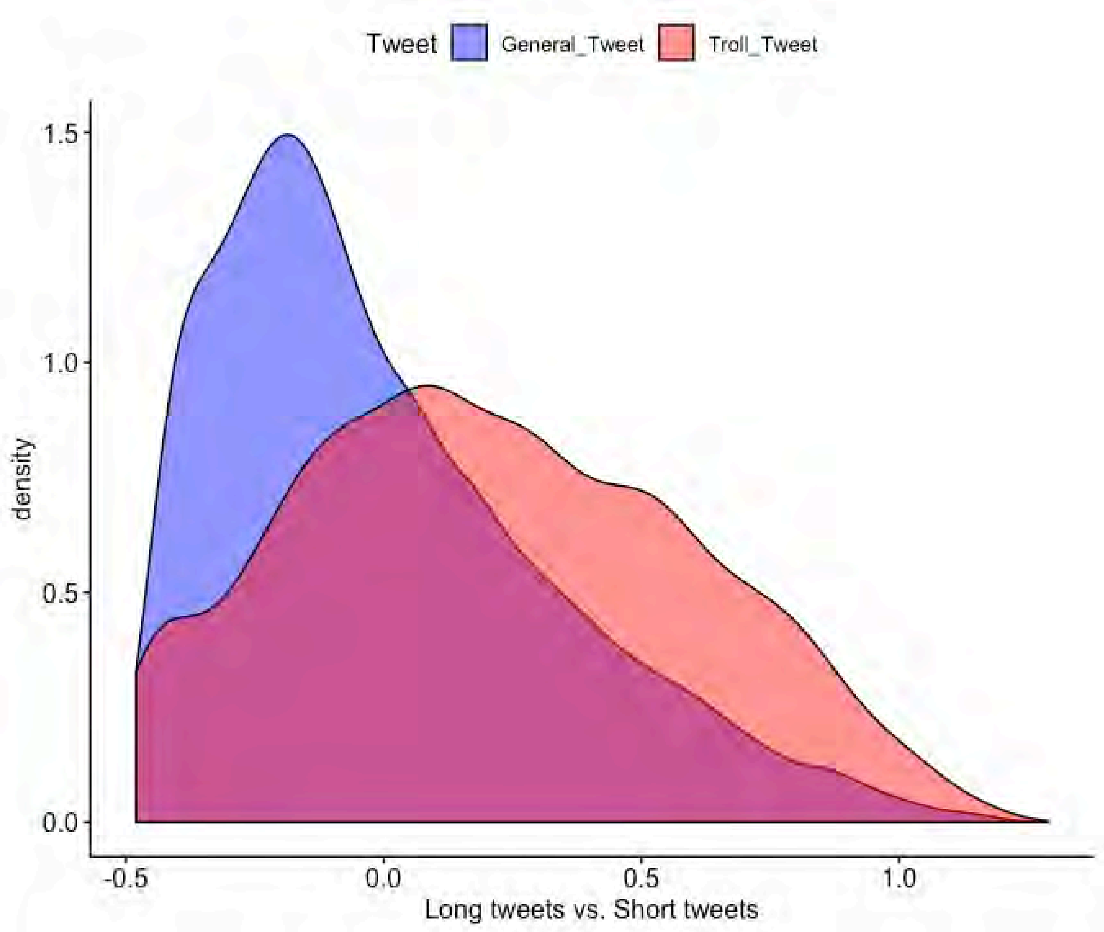
This dissertation so far has identified, described and compared the major patterns of linguistic variation of general English tweets and trolling tweets. It has been shown that trolling tweets and general tweets surprisingly share many of the same patterns of linguistic variation, and that they also have a distinct major linguistic repertoire. Whilst they may share some linguistic repertoires, it is not clear how trolling tweets compare to general English tweets with respect to the major patterns of linguistic variation of general English Twitter. Specifically, just because trolling tweets share many of the major dimensions of linguistic variation of general English Twitter, it does not mean that they are distributed similarly along them. Whilst this dissertation has described *how* trolling varies linguistically from one trolling post to the next, there has not been an investigation into *how* trolling varies linguistically in relation to general tweets. In other words, which linguistic repertoires of general tweets do trolling tweets align with the most? This chapter therefore compares trolling tweets according to the Dimensions of general English Twitter by projecting trolling tweets onto them.

To project trolling tweets onto the major dimensions of linguistic variation of general English Twitter, each trolling tweet was measured for the presence or absence of the 63 linguistic features (see Appendix 1) used in the MDA of general English Twitter, recording these results in a categorical data matrix. This data matrix was subsequently merged with the data matrix of general English tweets and this combined data matrix was analysed using MCA with general English tweets specified as active tweets, and the trolling tweets as supplementary (see section 4.7.2). The dimensions of linguistic variation of general English Twitter are not affected by this analysis. Rather, a coordinate is assigned to the trolling tweets revealing how associated it is to the dimensions. The overall positions of the trolling tweets with respect to each Dimension of general English Twitter is described below.

7.1. Long tweets vs. Short tweets

The first dimension of general English Twitter opposes long tweets with short tweets. Tweets assigned high positive coordinates on Dimension 1 tend to be longer in word tokens and have the presence of features, whereas tweets assigned high negative coordinates are shorter and tend to have the absence of features. Figure 6 is a density plot, which shows the distribution of each corpus with respect to General Twitter Dimension 1. Figure 6 shows that general tweets tend to be more clustered on the negative side of General Twitter Dimension 1, whereas trolling tweets are evenly spread along the dimension. In comparison to general tweets, Figure 6 shows that trolling tweets are on average assigned higher coordinates than trolling tweets on General Twitter Dimension 1, suggesting that trolling tweets tend to be longer than general tweets. This finding is supported in the mean tweet length of each corpus, where a tweet from the general Twitter corpus is on average 16 words long, and the mean tweet length of the Twitter trolling corpus is 24 words long.

Figure 6: Trolling tweets projected onto Dimension 1 of general English Twitter



7.2. Informational vs. Interactive

The second dimension of general English Twitter opposes tweets that have an informational broadcasting style with tweets that are more interactive. Tweets assigned high positive coordinates on Dimension 2 tend to be broadcasting to a large audience and comprised of numerous specific referents and carefully constructed and integrated information, characteristic of texts constructed when there is time to carefully plan and edit structures. Tweets assigned high negative coordinates tend to be directly conversing with other Twitter users and tend to employ an oral written kind of style, often encoding ones' stance.

Figure 7: Trolling tweets projected onto Dimension 2 of general English Twitter

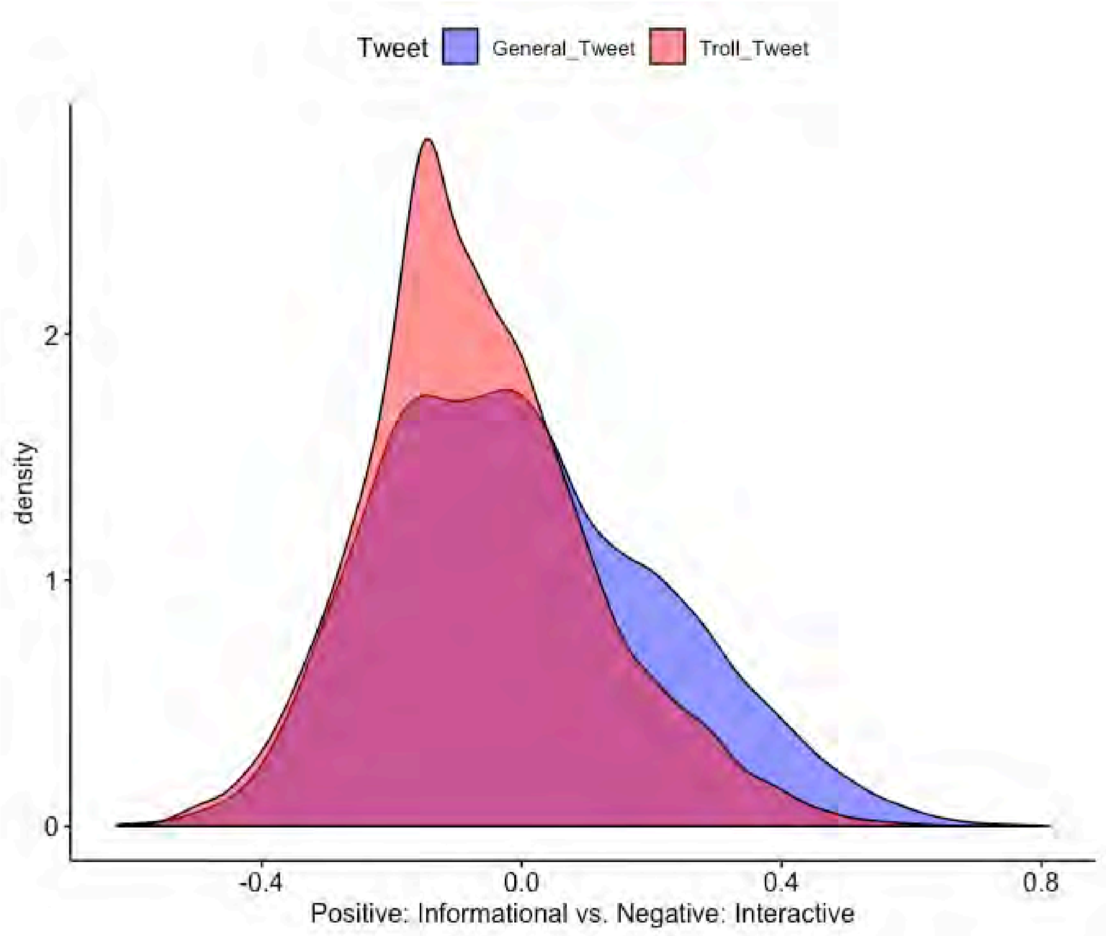


Figure 7 shows the distribution of each corpus with respect to the General Twitter Dimension 2. Figure 7 shows that there is considerable overlap between general tweets and trolling tweets with respect to Dimension 2, suggesting that trolling tweets and general tweets are very similarly distributed along this dimension. The density plot in Figure 7 shows that the proportion of trolling tweets that are situated on the negative side of Dimension 2 is slightly greater than general tweets, indicating that trolling tweets are only on average slightly more interactive than general tweets. Nevertheless, this difference is minimal and they are predominantly more similar than they are different with respect to this dimension. Overall, by projecting trolling tweets onto General Twitter Dimension 2, it has been shown that trolling tweets display a continuous amount of variation with respect to the degree of informational broadcasting/interactivity in a similar way to general tweets.

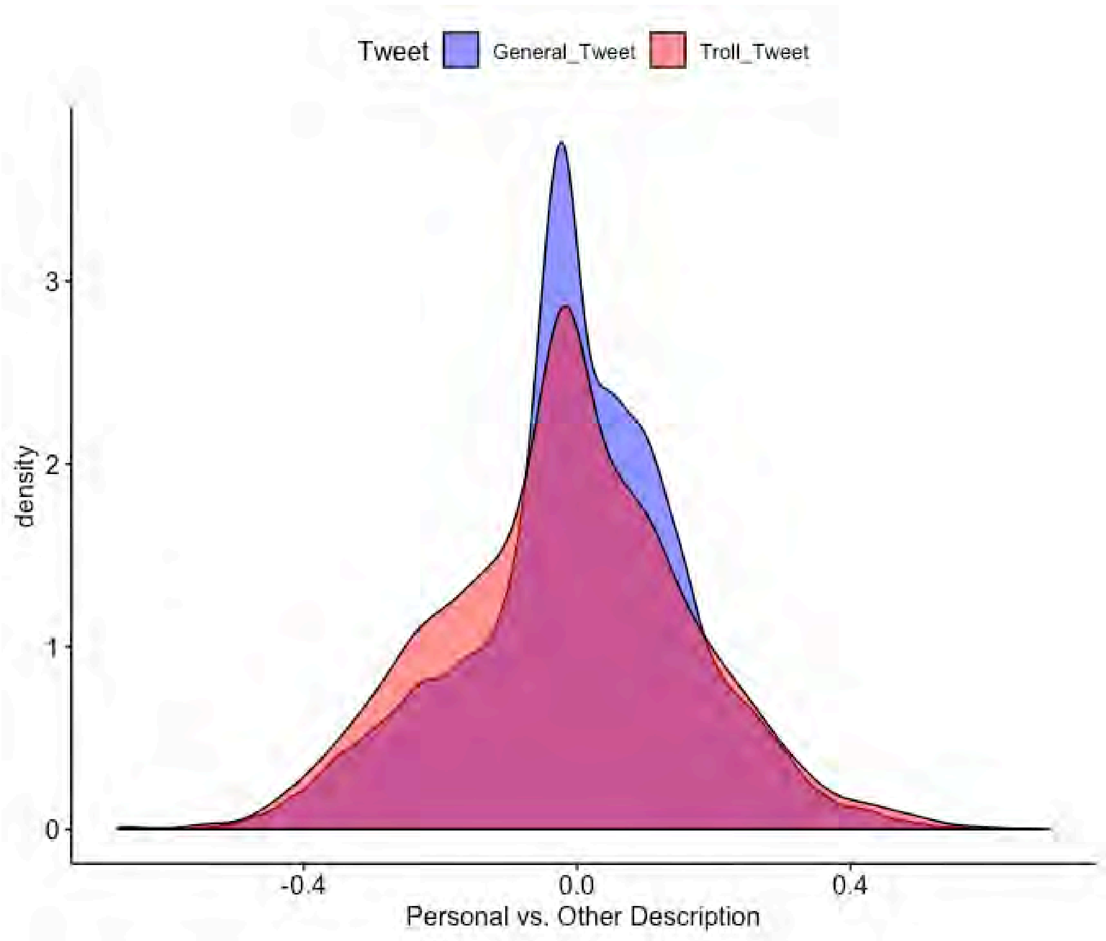
7.3. Personal vs. Other description

The third dimension of general English Twitter opposes tweets that have a highly personal and self-reporting style with tweets that are more focused on characterising and describing other entities. Tweets assigned high positive coordinates on Dimension 3 tend to be reporting on personal experiences and describe what the author is personally doing, thinking and feeling. Tweets assigned high negative coordinates on Dimension 3 tend to be describing, characterising and evaluating a subject that is not the self and is external to the conversation.

Figure 8 shows the distribution of each corpus with respect to General Twitter's Dimension 3. Figure 8 shows that the proportion of trolling tweets that are situated on the negative side of the dimension is slightly greater than general tweets, indicating that trolling tweets are on average slightly less personal than general tweets. This difference, however, is negligible. Importantly, Figure 8 shows there is considerable overlap between general tweets and trolling tweets with respect to Dimension 3, suggesting that trolling tweets and general tweets are very similarly distributed along this dimension.

Overall, this comparison indicates that trolling tweets display a continuous amount of variation with respect to the degree of personal/other description in a similar way to general tweets.

Figure 8: Trolling tweets projected onto Dimension 3 of general English Twitter



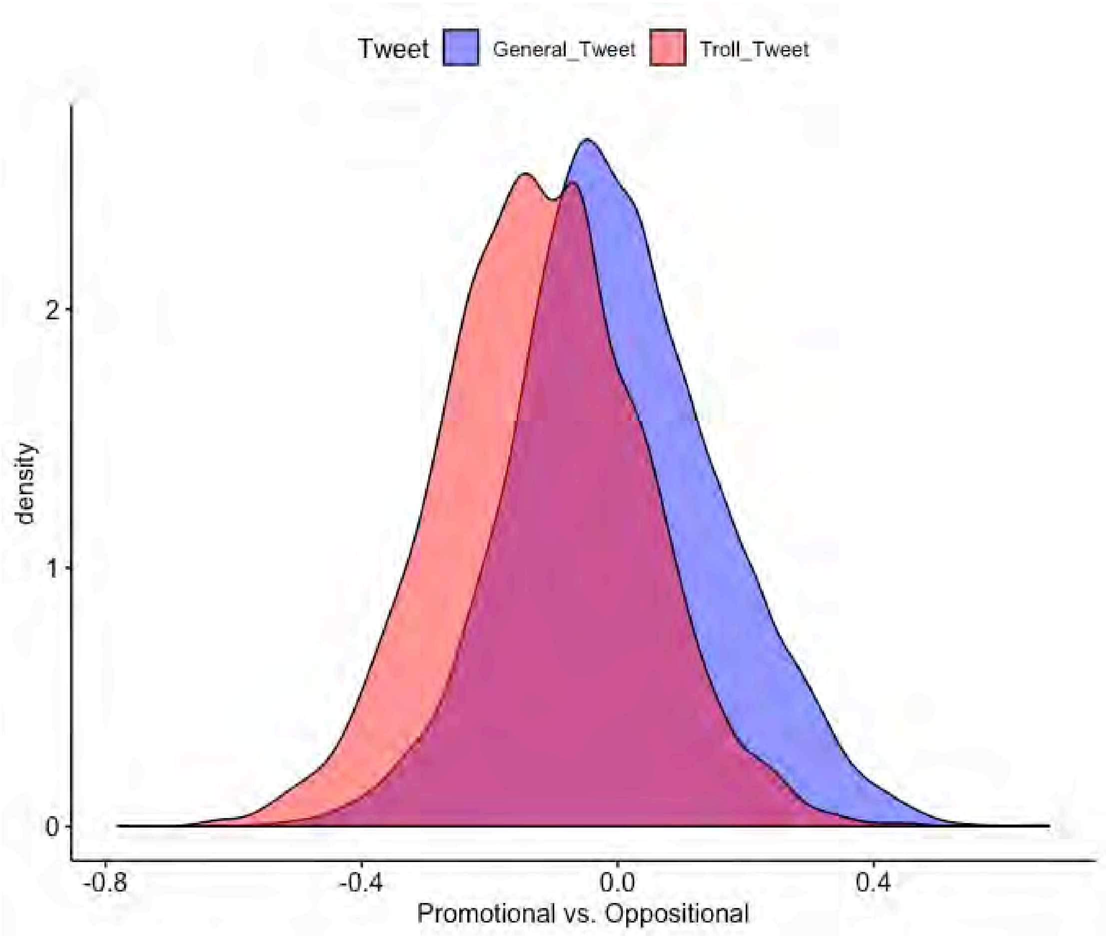
7.4. Promotional vs. Oppositional

The fourth dimension of general English Twitter opposes tweets that have a promotional style with tweets that are oppositional. Tweets assigned high positive coordinates on Dimension 4 tend to be attempting to gain attention and visibility by promoting and showing support for themselves and others. Tweets assigned high negative coordinates on Dimension 4 tend to be opposing, challenging, and disputing a particular stance, statement or ideology.

Figure 9 shows the distribution of each corpus with respect to the General Twitter Dimension 4. Figure 9 shows that more of the trolling tweets are associated with negative Dimension 4, whereas more of the general tweets are associated with positive Dimension 4, indicating that, on average, trolling tweets are more oppositional than general tweets. Figure 9 shows that trolling tweets that are most strongly associated with negative Dimension 4 have coordinates, which are considerably higher than any other general English tweet associated with negative Dimension 4. Additionally, Figure 9 shows that general tweets are much more associated with the promotional communicative style than trolling tweets. Although this is the greatest difference, Figure 9 shows that general tweets and trolling tweets largely overlap with respect to Dimension 4, indicating that trolling tweets and general tweets are similarly distributed along this dimension.

Overall, this comparison finds that trolling tweets display a continuous amount of variation with respect to the degree of promotion/opposition in a similar way to tweets.

Figure 9: Trolling tweets projected onto Dimension 4 of general English Twitter



7.5. Persuasive vs. Non-persuasive

The fifth dimension of general English Twitter opposes tweets that have a persuasive communicative style with tweets that are non-persuasive. Tweets assigned high positive coordinates on Dimension 5 tend to be attempting to bring something about in the future, often by demanding some form of action. Tweets assigned high negative coordinates on Dimension 5 tend to be reporting on past events and experiences and providing brief updates.

Figure 10: Trolling tweets projected onto Dimension 5 of general English Twitter

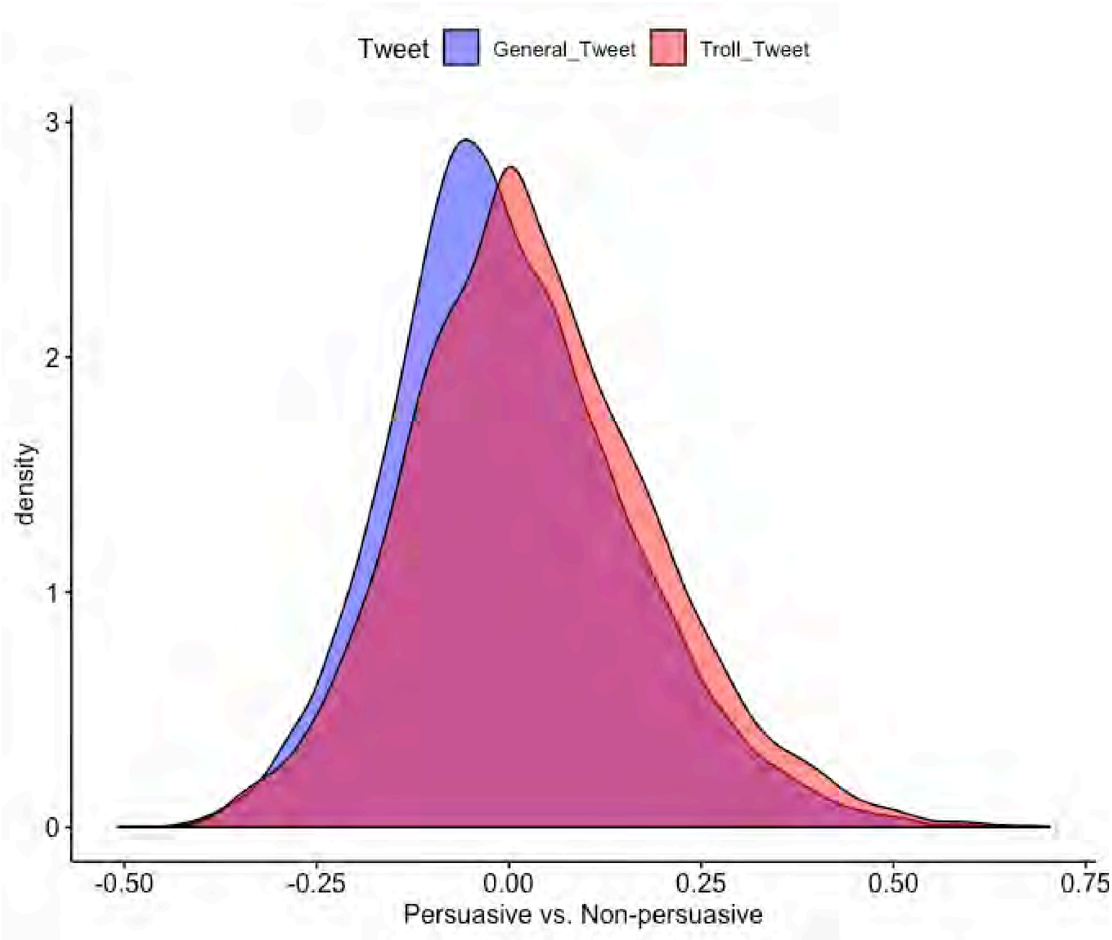


Figure 10 shows the distribution of each corpus with respect to the General Twitter Dimension 5. Figure 10 shows trolling tweets tend to be slightly more associated with positive Dimension 5, whereas general tweets are only slightly more associated with negative Dimension 5, suggesting that on average trolling tweets are more persuasive. This difference, however, is negligible, as there is considerable overlap between general tweets and trolling tweets with respect to Dimension 5, indicating that trolling tweets are distributed similar to general English tweets along this dimension. This is surprising considering that this pattern of linguistic variation was not a major dimension of linguistic variation in the MDA of Twitter trolling.

Table 48: Trolling tweets most strongly associated with positive and negative Dimension 5 of general English Twitter

	Trolling Tweet	D5 coord
1	@SLandinSoCal @realDonaldTrump #OperationMockingbird #JFKFiles Released files revealed at the CIA was controlling the media to generate propaganda. Don't believe what you hear on mainstream media.	0.6
2	OPINION: Hey, @NFL owners, tired of half-empty stadiums? This isn't hard. Tell your players to stand up for America https://t.co/yLWvXRXdwI	0.6
3	I hate gay people with the burning passion of a thousand suns. I wish I had bullied more of them in high school. The faggots deserve it	-0.5
4	Churchill was hated by his own party, opposition party, and press. Feared by King as reckless, and despised for his bluntness. But unlike Neville Chamberlain, he didn't retreat. We had a Chamberlain for 8 yrs; in @realDonaldTrump we have a Churchill.	-0.46

Because this dimension was not a major linguistic repertoire of trolling tweets in the MDA of Twitter trolling, some trolling tweets that are strongly associated with this dimension are presented in Table 48. Examples 1 and 2 are associated with a persuasive communicative style, whereas Examples 3 and 4 are associated with a non-persuasive, narrative communicative style. Example 1 is trying to persuade readers to disbelieve the media by telling them about Operation Mockingbird, which alleges that the CIA is controlling them. Example 2 is trying to persuade NFL owners to tell their players to stand up by suggesting that the reason why they have empty stadiums is because the athletes are kneeling during the national anthem in protest against racism, President Donald Trump, and police brutality. Both tweets are trying to convince readers about why something is the way it is, even though neither of the reasons is known to be the truth. Thus, these tweets may be trying to spread misinformation. Alternatively, Example 3 reports on past events in order to discriminate against and insult LGBTQ people. Example 4 also reports on past UK politicians in order to draw a comparison with Barack Obama and Donald Trump to negatively present and insult Obama and positively present Donald Trump. Overall, the examples in Table 48 show how these communicative styles can be employed for a variety of communicative goals.

7.6. Discussion

This dissertation sought to provide a thorough linguistic description of Twitter trolling not only with respect to its major linguistic repertoires and communicative functions, but also in relation to general Twitter posts. Having identified and described the major patterns of linguistic variation of general tweets (Chapter 5) and trolling tweets (Chapter 6), this chapter sought to identify and describe how trolling tweets vary in relation to the major patterns of linguistic variation of general tweets by projecting trolling tweets onto the Dimensions of general English Twitter.

The results of this comparison show on average trolling tweets are longer than general tweets. Additionally, the results indicate that trolling tweets are on average slightly more interactive, oppositional and persuasive, and slightly less personal than general tweets. Nevertheless, these mean results hide the remarkable and significant overlap between trolling tweets and general tweets with respect to *all* of the dimensions of linguistic variation. The results show that trolling tweets not only share the major linguistic repertoires of general tweets, but that trolling tweets vary along these patterns of linguistic variation in much the same way as general tweets. Figures 6 – 10 show that trolling tweets are distributed similarly to general tweets along the dimensions of general English Twitter, indicating that trolling tweets and general tweets are far more similar than they are different. Although it is not clear just how many of the general tweets are trolling tweets, even if they constituted 30% of the general Twitter corpus, the amount of overlap would still be significant and remarkable.

This considerable overlap across all dimensions may be partially explained by the theory that some trolls are trying to make their posts seem genuine in order to provoke a reaction. Some trolls need to disguise their trolling intentions because otherwise they may be less likely to provoke a reaction (Donath, 1999; Hardaker, 2010). One way for trolls to appear genuine may be to "blend in" to the general practices of Twitter - essentially to assimilate to general tweets. Cruz et al. (2018) suggest that assimilation is essential for trolling to be successful, as a way to deceive one's audience. They suggest that assimilation is similar to learning in that the troll needs to be aware of the other users and their behaviours. The troll also needs to take note of how the users are perceived and recognised within the community. Once the troll has this knowledge, they need to then imitate. The considerable overlap between trolling tweets and general tweets with respect to the Dimensions of linguistic variation of general Twitter may be because trolls have learned the major repertoires of tweeting and are emulating these repertoires in a process of assimilation in order to disguise their trolling intentions. Importantly, if this is the case, the results show that trolling tweets are exceptionally good at assimilating. Whilst some trolls may indeed be drawing on the major linguistic repertoires of general tweets for the purpose of assimilation, there is no guarantee that all trolling tweets in the corpus are aiming to provoke a response and are trying to deceive their audience. Thus, this may only offer partial explanation for the overlap.

Another reason for such overlap may be because trolls are trying to adopt and exploit the communicative practices of Twitter in a process of détournement (Debord, 1967), whereby trolls are culture jamming - posting like other Twitter users in order to mock them, so that when other Twitter users critique trolls for their behaviour, then it turns out that the Twitter users are actually unknowingly critiquing themselves (Phillips, 2013; 2016). This raises the question about whether trolling is all that bad. Whether it is intentional or not, by employing the same linguistic repertoires, the trolls position their posts on the same level as their victims and general Twitter and point the finger at general Twitter. It can be argued that trolling in this way serves to remind Twitter that everyone is entitled to their opinion, and that critique, debate and diversity of views are important for a healthy society. This is not to suggest that trolling and spreading misinformation is good. The abusive trolling in Dimension 5 of Twitter trolling and fake advice that causes individuals to damage their property are just two examples of questionable behaviour. The point here is to suggest that if Twitter trolling is détourning Twitter practices, then trolling may just be a symptom of a larger problem of Twitter more generally. This point will be developed in the following chapter.

Whilst the notion of détournement can explain the overlap between trolling tweets and general tweets, the theory does not hold for all trolling. Not all instances of trolling are purposefully tweeting in the same way as general tweets in order to mock general twitter users. Spreading misinformation can be for the purpose of driving conspiracy, among others. Thus, détournement explains part of the data but not the whole.

Another reason for the extraordinary amount of overlap may be because Twitter is so conducive to trolling. Trolling is often aimed at humiliating their victims through a variety of different strategies, including making them believe fake information, provoking a response, insulting them and so on. Although some accounts are anonymous, much of Twitter is related to identity construction and presenting one's best and most successful self. In a context that is so tied to one's identity, trolling becomes so much more problematic, as it is a direct attack on one's carefully crafted identity. When considered from the theory that trolling is aimed at laughing at other people's expense, Twitter, which involves millions of people encoding their stance, reporting on their daily lives, promoting themselves and presenting themselves positively, is arguably very conducive to trolling, as each tweet provides trolls with fuel for accumulating lulz. Moreover, Twitter is a highly public platform, which enables trolls to observe their victims easily and understand what is likely to antagonise them.

Additionally, Twitter is also very conducive to trolling due to its technological affordances. Research has noted that the technological affordances of Twitter facilitate negative and transgressive behaviours such as, impulsive and uncivil discourse (Ott, 2017), public shaming (Nicotra, 2016), and the spread of misinformation (Dumenco, 2011). For example, the case of Justine Sacco illustrates exactly how Twitter can be prone to a lack of forethought and uncivil behaviour, as described by Ott (2017), as not only did she send a careless, racist tweet (Pilkington, 2013), but also she was publicly shamed for doing so. Sacco’s tweet was retweeted thousands of times with various comments and numerous Twitter users talking about it. Specifically, the hashtag ‘#HasJustineLandedYet’ was used so many times that it became a trending topic. As a result of these technological affordances, which enabled the tweet to go viral (e.g. non-reciprocity of following, @mentioning, retweeting, hashtagging and trending), Justine Sacco was publicly shamed as people criticised Sacco, labelled her “a racist”, and wished AIDs upon her. Importantly, Sacco only had fewer than 200 followers; however, her tweet and the reactions to her tweet were accessible to people all over the world via these technological affordances. Although Sacco arguably deserved some retribution for her racist tweet, the extent of the retribution that she received as a result of the technological affordances of Twitter was damaging. Based on the results of this chapter, it may be argued that the linguistic repertoires of general Twitter may also be conducive for trolling. For example, the ability to interact with any Twitter user enables trolls to target others and provoke a response from them. The ability to broadcast one’s message to a larger audience enables trolls to spread hyperbolic or fake news. Moreover, the results demonstrate that trolls can exploit the common repertoires of general tweets for the purpose of trolling. Trolls can encode a personal stance and experience, or characterise and describe others (Dimension 3) in order to insult or spread misinformation. The same can be said for General Twitter’s Dimension 4, where trolls can oppose or show support for particular ideologies, people or statements to mock, silence, insult or provoke others. Finally, a persuasive or non-persuasive and narrative communicative style can also be employed by trolls in order to convince individuals to believe certain theories that have not been proven yet and spread fake news, or by reporting on past events that are negatively marked and insult and mock particular people.

Overall, each theory may explain part of the overlap between trolling tweets and general tweets. However, individual trolls may have different communicative goals and ends, meaning that there is no single theory that is able to account for all of the data. The findings of this dissertation provide empirical evidence about the major communicative functions of general tweets and trolling tweets and the results show how these communicative functions can be exploited by trolls for their individual trollish ends. Thus, it is argued that trolls may be drawing on the major communicative functions and patterns of linguistic variation of General Twitter in order to pursue their individual interests.

7.7. Conclusion

Importantly, even though trolling tweets are situated in the context of general Twitter and are a type of tweet, they are not predisposed to have the same major linguistic repertoires or vary along these dimensions of linguistic variation in the same way. Whilst there are slight differences in the overall tendencies of the trolling tweets with respect to the major patterns of linguistic variation, these differences are only observable when aggregating them altogether. Thus, the degree of similarity between trolling tweets and general tweets is extraordinary and has particular implications for researchers looking at detecting it automatically. The results show that both groups display a continuous range of linguistic variation, and currently the tools used to detect them are not based on such intra-variation. Importantly, the results also expose the fact that a binary distinction between that which is trolling and that which is not with respect to these major patterns of linguistic variation may not be possible, which problematises the notion that it can be detected automatically. Moreover, the results demonstrate that by believing that these two groups can be distinguished is perhaps an oversimplification, especially given the deceptive nature of some trolling instances and/or the possibility that some trolls could be détourning general tweeting practices.

8. Discussion

This chapter discusses the significance of the results and the methodological contributions of the dissertation. It begins with a summary of the results and then a discussion of the significance of these results in light of previous research and the research questions. Subsequently, this chapter examines the methodological contributions of this dissertation. Finally, the limitations of the dissertation are acknowledged and future research directions are introduced.

8.1. Summary of Results

Before this dissertation, research had established that trolling is a diverse and deceptive phenomenon that varies not only with respect to its different behaviours, types, motivations and guises, but also in terms of how it is perceived. Trolling means different things for different people in different contexts. For anyone interested in language variation, this raises the questions about whether such variation is reflected in the language of trolling, and also how the language of trolling varies in relation to general social media posts. Despite this, an investigation into understanding how the language of trolling varies was yet to be conducted with most research focusing on how to detect it automatically from non-trolling social media posts. This body of research has developed classifiers trained on language patterns and other features that can detect trolling with high accuracy, demonstrating that trolling and non-trolling posts are largely different. The thing missing in these studies was a linguistic explanation about *why* the classifiers worked and *why* certain features were perhaps more predictive than others.

In light of this, this dissertation sought to provide the first quantitative linguistic description of Twitter trolling in terms of its major patterns of linguistic variation, as well as with respect to the major patterns of linguistic variation of general English tweets. Collecting trolling posts, however, can be particularly difficult, especially considering that trolling can vary so much and that it can be a deceptive phenomenon. To avoid imposing the researcher's opinion on what constituted trolling and to attempt to get a more varied sample, this dissertation collected trolling posts by using other people's perceptions of trolling. Tweets accused to be trolling were collected, resulting in a corpus of 4,182 trolling tweets, which is, to my knowledge, the largest collection of perceived trolling tweets to date.

In addition to a Twitter trolling corpus, a general Twitter corpus was collected from the random 1% stream of public tweets using the public API to serve as a baseline of Twitter for comparison. There is no doubt that this general Twitter corpus contains trolling posts, spam, and bot tweets. Unlike NLP research, this dissertation was not aimed at detecting trolling posts from non-trolling posts, but rather it was aimed at understanding how trolling tweets vary in relation to general tweets. General tweets in this case are not non-trolling posts. Rather trolling tweets have become an integral part of Twitter, and thus a general Twitter corpus that does not contain spam and trolling tweets fails to represent general Twitter. Nevertheless, the amount of trolling tweets in this corpus is not known.

To identify the major patterns of linguistic variation across these two corpora of tweets, this dissertation introduced a modified version of MDA, which is based on a Multiple Correspondence Analysis (MCA) of the occurrence of linguistic features in each tweet. This MCA-for-MDA was applied separately to each corpus based on the linguistic features that occurred in more than 5% of the tweets in each corpus. Overall, 5 major dimensions of linguistic variation of general English Twitter, and 5 major dimensions of linguistic variation of Twitter trolling were identified and described, as these were readily interpretable. Given that the two corpora were collected with different constraints and that different (although largely similar) feature sets were used in each analysis, it was expected that each analysis would return several distinct dimensions of linguistic variation. Remarkably, however, it was discovered that the first four dimensions of linguistic variation across both corpora were exceptionally correlated, indicating that these dimensions of linguistic variation were perhaps shared. From the interpretation of the dimensions and the tweets most associated with them, it was shown that the exceptional correlations were indeed realised the same in the two sets of tweets.

The first shared dimension of linguistic variation opposes long tweets with short tweets. This dimension shows that longer tweets tend to have the presence of features, whereas shorter tweets tend to have the absence of features. As the first major dimension of linguistic variation, it indicates that the strongest influence on the presence or absence of linguistic features across the trolling and general tweets is the length of the tweet, as the more words a tweet has the more likely it is to have a range of linguistic features.

This dimension is distinct from other MDAs, which measure the relative frequencies of features, as opposed to their presence or absence, because the analysis of relative frequencies of features enable text length to be controlled for, whilst the occurrence of features does not. Apart from a slight correlation to Dimension 2, tweet length is only strongly correlated to Dimension 1. All other dimensions show no correlation, suggesting that tweet length has been largely controlled for in the first dimension. This pattern is crucial in enabling the interpretation of the rest of the linguistic co-occurrence patterns in a similar way to standard MDA because if Dimension 1 was not associated with tweet length, the other dimensions could have been confounded by tweet length, which might have made the interpretations less coherent. Given its importance to the rest of the analysis, rather than ignore it, it might be viewed as a “pre-dimension”, as tweet length is not a communicative function.

The second shared dimension of linguistic variation opposes informational broadcasting tweets with interactive tweets. This dimension indicates that the next most important distinction of tweets opposes tweets that are specifically interacting with other Twitter users with tweets that are publicly broadcasting their message. Although the technology now affords these two communicative patterns (Yaqub et al., 2017), interaction was not the purpose of Twitter, rather tweets were, by default, designed to be broadcast to the public stream of tweets and appear in the timelines of one’s followers to update them on one’s activities (Zappavigna, 2015). Users nevertheless had the need to communicate with each other and began interacting with each other on Twitter by using the ‘@’ symbol followed by the username, which derived from Internet Relay Chat (Werry, 1996). As a result of its prominence on the platform, Twitter incorporated @mentioning into its design (Weller et al., 2014; see Halavais, 2014). This dimension therefore represents the influence and symbiotic relationship of the creators, the users and the technology. As the second most important dimension of linguistic variation, it demonstrates that interaction and the public broadcast of information are important communicative functions across Twitter and Twitter trolling. This pattern of linguistic variation has been observed across nearly all MDA studies (Biber, 2014) and indicates that trolling tweets and general tweets can employ a more oral spoken kind of style, or that they can employ a more literate and carefully planned written form.

The third shared dimension opposes tweets with a personal focus with tweets that describe and characterise things other than the self. This dimension shows that tweets vary in regards to whether they are about the self or whether they are about others. Blogs also vary in this way (Grieve et al. 2010), which offers linguistic support to common characterisations that Twitter is a microblogging service. Tweets about the self tend to concern a variety of different personal experiences, feelings, thoughts and ideas, and can be likened to overt expressions of identity (Bucholtz and Hall, 2005). Tweets about others tend to involve describing and characterising entities, which encodes one’s stance in relation to them, which can be a more covert identity expression. As the third major dimension of linguistic variation, it suggests that expressions of identity, especially encoding one’s stance and describing personal experience are important communicative goals for general Twitter users and Twitter trolls.

The fourth shared dimension opposes promotional tweets with oppositional tweets. This dimension indicates that tweets vary in regards to whether they align or not with particular positions. Tweets that are promotional are associated with demonstrating support and alignment and gaining attention, whilst tweets that are oppositional are associated with challenging a position. This dimension is suggestive of a diversity of users, each with their own interests and ideological positions. Moreover, as the fourth major dimension of linguistic variation, it marks that positioning oneself as affiliated or disaffiliated is an important communicative goal for Twitter users and Twitter trolls, which is another resource for identity expression (Bucholtz and Hall, 2005). This dimension also shows that Twitter affords the ability to reach broad audiences and gain attention, which is often exploited by Twitter users for promotional purposes.

The two corpora differ on their fifth major dimension of linguistic variation. Dimension 5 of general Twitter opposes tweets that are persuasive with tweets that are non-persuasive and narrate on personal past events. This dimension indicates that general tweets vary according to whether they aim to bring something about in the future or whether they are reporting on past events. This dimension suggests that Twitter can be used to demand and gain something. Moreover, it suggests that Twitter can be used like a public personal diary, which aligns with its original intended use for posting short updates for the consumption of friends (MacArthur, 2019). As the fifth major dimension of linguistic variation of general Twitter, it shows that persuasion and personal narration are important communicative goals for Twitter users, which suggest that Twitter is used commercially, as users not only engage in persuasive techniques for economic or non-economic gain, but that they also produce tweets for the consumption of others.

Dimension 5 of Twitter trolling opposes uncivil tweets with civil tweets. This dimension indicates that trolling tweets vary according to whether they are abusive and antagonistic or whether they are more respectful and conciliatory in their argumentative style. This dimension indicates that argumentation is an important communicative goal for Twitter trolling, which aligns with the notion that trolls are oppositional and aim to provoke a response (Hardaker, 2010; Phillips, 2016), as arguments are common when there is an opinion conflict. Although it is not clear whether general tweets also vary along an uncivil/civil communicative style, as the fifth major dimension of linguistic variation of Twitter trolling, the results of the analysis show that this particular communicative style is far more prominent in trolling tweets than general tweets.

These dimensions of linguistic variation show that there is a great deal of linguistic variation across general Twitter and Twitter trolling. The dimensions of linguistic variation of general Twitter have been explained by the technological affordances and promoted uses of Twitter, the genre of microblogging, the commercial use of Twitter (for economic and non-economic gain), and according to the diversity of users and their major communicative goals (see Chapter 5). Whilst the dimensions of general Twitter and these explanations may seem natural and recognisable for anyone experienced with Twitter, many of the dimensions that are shared with trolling are contrary to common knowledge and descriptions of trolling. For example, a personal and promotional communicative style, as described on the positive poles of Dimension 3 and 4 respectively, run counter to the notion that trolls need to remain detached and emotionally divested (Phillips, 2011), and that trolls are intentionally oppositional (Phillips, 2016), as opposed to supportive. Notably, even the civil argumentative style described on the negative pole of Dimension 5 of Twitter trolling is also at variance to common descriptions of trolling, as trolls are frequently described as being uncivil and having a hostile and abusive argumentative style (Synnott et al., 2017).

The unexpectedness of these communicative behaviours may work in the trolls' favour, especially considering the fact that some trolls aim to provoke a response. Because such communicative styles have not been described as common repertoires of trolls, the likeness of the trolling tweets to general tweets may convince other Twitter users that the troll is a genuine poster, which may be more likely to provoke a response from their victim. Thus, whilst the linguistic repertoires may seem counterintuitive to descriptions of trolling, they could be purposeful linguistic realisations of the notion that some trolls aim to deceive their victim into thinking they are posting genuinely so that they are more successful in provoking a response (Donath, 1999; Hardaker, 2010).

Whilst the first four dimensions of trolling tweets were exceptionally correlated to general tweets, demonstrating that they share many of the major linguistic repertoires, the correlations do not suggest that trolling tweets vary along the dimensions in the same way as general tweets. The majority of trolling tweets, for instance, could cluster on particular sides of the dimensions of general tweets. For example, given the notion that trolling is intentionally oppositional (Phillips, 2016), it might be expected that the majority of trolling tweets would be associated with the oppositional communicative style (negative Dimension 4 of general Twitter), whilst general tweets might be largely associated with the promotional communicative style. Thus, in Chapter 7, this dissertation sought to compare the distributions of trolling tweets in relation to general tweets.

To compare their distribution, trolling tweets were projected onto general Twitter's dimensions of linguistic variation. On aggregate, slight differences between general tweets and trolling tweets were discovered. Trolling tweets were found to be longer, more interactive, more oppositional, more persuasive and less personal than general tweets. These slight differences partially support previous research that has found that trolls are marked by their hostility, emotional detachment, dialogical provocativeness (i.e. trying to get a response), and their often-convincing dissemination of exaggerated or false information (Phillips, 2016; Phillips, 2011; Hardaker, 2010; NetLingo, 1995-2015). Nevertheless, these slight overall differences hide the overarching similarity between trolling tweets and general tweets in their distribution along the dimensions of linguistic variation. Specifically, trolling tweets and general tweets overlap considerably on each dimension, showing that trolling tweets are far more similar to general tweets than they are different, not only in terms of the range of communicative styles found across them, but also in terms of their distribution and use of these styles.

8.2. Significance of the Results

The first research question sought to identify and describe the major patterns of linguistic variation across general Twitter. Although this analysis of general Twitter was predominantly intended to serve as a baseline for the comparison of trolling tweets, it also serves in and of itself as a major contribution to the field. Whilst research has previously described different purposes of tweets (e.g. Lee, 2011), and the overall association of groups of tweets with respect to the range of linguistic distributions of other language varieties (e.g. Sardinha, 2014; 2018; Friginal, Waugh and Titak, 2018; Coats, 2016; Passonneau et al. 2014), this is the first study of its kind to describe the full range of linguistic distributions across English Twitter and examine patterns of linguistic variation between individual tweets, as opposed to concatenated groups of tweets.

By focusing on Twitter specifically, as opposed to Twitter and other varieties of language, this dissertation finds dimensions of linguistic variation that are not influenced by the inclusion of other registers, but are instead specific to this corpus of tweets. Previous research has described that Twitter is used for a variety of purposes, such as for disseminating information on major news events (Zappavigna, 2018), for personal expression and interaction (Honeycutt and Herring, 2009), as well as for recording their thoughts and narrating their everyday activities (Weller et al., 2014). The results of this dissertation provide linguistic support to these findings and also contribute additional purposes based on the major patterns of linguistic variation. Twitter is used for interacting, broadcasting, personal expression, and describing others. It is also used for promoting the self and others and showing support and opposing particular statements and entities. Finally, the results show that Twitter is used to persuade others and obtain some form of capital (economic and non-economic) and narrating personal events.

The results demonstrate that a huge amount of variation exists between individual tweets, suggesting that the factor scores of concatenated groups of tweets in other MDA studies may hide the significant differences across individual tweets. For example, Sardinha (2014) found that groups of tweets were highly associated to the involved and interactive pole of Biber's (1988) Dimension 1. Additionally, Sardinha (2018) observed that groups of tweets were associated to the involved and interactive pole of Dimension 1 of Internet registers. The findings from the present dissertation show that tweets can also be informational, indicating that it is important to analyse them individually.

Friginal, Waugh and Titak (2018) found that tweets generally were informational along the dimensions of Internet registers described in Titak and Roberson (2013). They also looked at groups of tweets defined by topic and found that the tweets that were interactive and involved tended to be those on the topic of sports and entertainment, whereas tweets on personal topics, the weather, politics, and business were more informational. Although Friginal, Waugh and Titak (2018) grouped the tweets by topic, which enabled tweets to fall along the continuum of interactive and informational, this still limits the results in identifying the most general pattern of the topics, rather than the specific patterning of individual tweets. The findings from the present dissertation demonstrate how a tweet with a political focus (Example 18, Chapter 5) can be highly interactive, whereas a tweet on the topic of entertainment (Example 11, Chapter 5) can be exceptionally informational, which is far more specific at the individual tweet-level than the findings of Friginal, Waugh and Titak (2018).

The results of the MDA of general Twitter have particular significance for tools and linguistic tasks, which are influenced by or can be enriched with information on register and communicative purpose, such as authorship analysis. This dissertation demonstrates that language varies on Twitter for particular communicative functions. Authorship analysis involves finding consistent language patterns in one author that are distinctive to another author’s consistent language patterns. Nevertheless, register can influence the particular language patterns. Thus, this method enables register to be isolated so that author style can perhaps be identified.

The results of general Twitter suggest that the design of the technology has influenced some of the major linguistic repertoires. In particular, Dimension 4 opposes tweets that are promotional with tweets that are oppositional. Both of these communicative styles have been described as being influenced by the need to gain attention and more followers, which has been incentivised by the follower size being in big and bold (Page, 2012; Dorsey, 2019; Anderson, 2019). As the fourth major linguistic repertoire, it suggests that particular design features can have quite a big influence on the language use of Twitter. Research notes that oppositional and rational-critical discourse (Dahlberg, 2001) is important for a healthy society. Additionally, the linguistic resources associated with promotional tweets have been important for silenced individuals, marginalised groups and for organising protests and political activism because they are useful in garnering attention and showing support for particular ideas and positions. Whilst this research indicates that these communicative styles can certainly be essential for particular communicative goals, if they are exploited for malevolent, egotistic and narcissistic ends, such as for gaining more followers, to humiliate someone, or to gain some form of status, then arguably they become more problematic. Thus, it is important that social media creators consider the influence of particular design features, as the findings from the research can be used to suggest that they have influenced the language.

The second research question sought to identify and describe the major patterns of linguistic variation across Twitter trolling. Given that *trolling* is used to label a variety of behaviours, this dissertation sought to examine whether such behavioural distinctions were reflected in language differences. The results demonstrate that, whilst a huge amount of variation occurs across trolling tweets, the first four major linguistic repertoires are the same as general English Twitter, indicating that trolls are drawing on the resources of general tweets in order to troll.

Whilst the results of the analysis of Twitter trolling challenge previous descriptions of trolling that emphasise its oppositional, hostile and emotionally detached nature, the remarkable similarity between general tweets and trolling tweets in their major linguistic repertoires simultaneously partially informs the theory that trolling is a deceptive practice (Donath, 1999). Some trolls aim to provoke a response and will say and do anything to get a rise out of someone, even if they do not agree with it themselves (Phillips, 2011). Because some trolls’ main purpose is to obtain the reaction, a common piece of advice is to starve the trolls of the reaction they crucially crave by not engaging with them or responding to their comments. This means that the trolls risk being unsuccessful in gaining a response if their victim suddenly notices that they are trolling. Thus, a necessary task for this kind of troll is to disguise their trolling intentions and make their provocative comments appear genuine (Donath, 1999; Hardaker, 2010).

One way in which trolls can appear genuine is to employ communicative behaviours that are unexpected and/or similar to ordinary Twitter users in a process that has been labelled ‘assimilation’ (Cruz et al., 2018). Assimilation involves learning the behaviours of the online community and how the users are perceived and recognised within the community (Cruz et al., 2018). Once established, the trolls can imitate and may appear genuine (Cruz et al. 2018). Trolls may be drawing on the same major linguistic repertoires of general tweets in a process of assimilation in order to deceive their audience and appear genuine, so that they are more likely to provoke a response. Moreover, they may employ characteristically unexpected communicative styles like a civil style also as a way to disguise their trolling intentions and provoke a reaction.

Finally, the fact that Dimension 5 does not align with general tweets may also be explained in light of deception and *leakage*. In interpersonal deception theory (Buller and Burgoon, 1996), leakage refers to the notion that deception is cognitively demanding and that sometimes the underlying intentions (i.e. the truth) may leak out and be revealed. Leakage is predominantly understood as non-verbal signals, such as saying yes whilst shaking one’s head. However, in this case, leakage could be the use of one communicative style over the other. The results may therefore suggest that we can learn a lot from trolling with respect to strategies of deception and the notion of leakage.

The third research question sought to compare trolling tweets to general English tweets to gain a complete linguistic understanding about how trolling tweets vary in relation to general tweets. The results of this comparison not only demonstrated that trolling tweets draw on many of the same linguistic repertoires of general tweets, but it also showed that the trolling tweets are distributed similarly along General Twitter's dimensions of linguistic variation. These results show that trolling tweets are remarkably more similar to general tweets than they are different.

The remarkable similarity between trolling tweets and general tweets with respect to general Twitter's dimensions of linguistic variation has particular implications for NLP researchers that aim at detecting it automatically. The results show that both groups of tweets display a continuous range of linguistic variation, and currently the tools used to detect them are not based on such intra-variation. The results also expose that a binary distinction between trolling tweets and general tweets with respect to these major patterns of linguistic variation may not be possible, which problematises the notion that it can be detected automatically using such grammatical patterns. This is not to say that it cannot be detected using other features (linguistic and non-linguistic). The state-of-the-art NLP models are able to distinguish some kinds of trolling posts from non-trolling posts but they are not able to distinguish both groups with complete accuracy (see section 2.3.4). The findings from the present dissertation may therefore inform future method design as they demonstrate that lexico-grammatical variation is not very informative in a classification task.

Additionally, the results demonstrate that by believing that these two groups can be distinguished is perhaps an oversimplification, especially given the possibility that some trolls are deceptive and/or the possibility that some trolls could be détourning general tweeting practices. The results show that the boundaries between the major linguistic repertoires of trolling and general tweets are blurred and fuzzy, and that context and the audience play a huge role in distinguishing trolling. For example, saying “I love Trump - he is the best president the United States has ever had” in one community can be normal and accepted; whilst in another context it is deviant and provocative. At the same time “You fucking dickhead” can be completely harmless amongst friends. These findings suggest that dimensions of linguistic variation may not be informative in distinguishing trolling. This has implications for forensic linguistics, as sometimes linguists may be brought in to the legal context to determine the offensiveness of a particular social media post and/or whether it was trolling or not. The results suggest that particular lexico-grammatical patterns do not distinguish trolling from general social media posts. Suggestions that these language features clearly show that something is or is not trolling must therefore be taken with caution.

Although the major linguistic repertoires of general tweets are not able to distinguish trolling from general social media posts, the fifth major linguistic repertoire of trolls identified in this dissertation may be important for NLP researchers. Whilst it is not clear whether general tweets also employ Dimension 5 of Twitter trolling, the differences between the fifth major dimensions across the two corpora may offer some linguistic explanations for NLP researchers as to why certain linguistic features worked in their classifiers. The results show that profanity patterns differently in general tweets and trolling tweets. Positive Dimension 5 of Twitter trolling shows that profanity co-occurs most often in trolling tweets with question features (WH-words, question marks), features associated with making demands (imperatives), interactive and targeted features (second person pronouns, initial mentioning), features for emphasis and gaining attention (hashtags, URLs, capitalisation), features for characterising and describing (stative forms) and many forms associated with a present tense orientation. In the context of trolling tweets, these linguistic features co-occur with profanity for the purpose of abusing and being uncivil. Alternatively, negative Dimension 5 of general Twitter shows that profanity co-occurs most often in general tweets with personal narrative features, often for the purpose of evaluating a personal experience or event in a non-targeted and non-abusive way. Given that most NLP studies investigating trolling collect trolling posts by searching for profanity and slurs (Golbeck et al. 2017), these different linguistic co-occurrence patterns could be implemented into classifiers to test if they have predictive power. Moreover, in previous studies that have incorporated different features into their classifiers, the results of this dissertation may be used to explain why the features worked, especially if they align with many of the features associated with an uncivil style.

On a similar note, the results of this dissertation echo Ajayi (2018), Clarke and Grieve (2017), and Waseem et al. (2018) demonstrating that trolling and abuse does not have to contain profanity, and thus future NLP research should focus on other more subtle forms of trolling and abusive language. The way in which this dissertation collected trolling posts could also be applied for other kinds of abusive language, such as racist and sexist tweets, which may enable more subtle forms of abuse to be accounted for in classifiers.

8.3. Methodological Contributions

8.3.1. The Short Text Version of MDA

This thesis introduces a modified version of Biber's (1988) MDA, which is able to identify patterns of functional linguistic variation between individual tweets. MDA has traditionally been used for identifying and analysing the frequent patterns of linguistic co-occurrence across a corpus of texts. The approach is based on a factor analysis of the relative frequencies of grammatical forms across the text samples in the corpus. Because the main objects of analysis are the relative frequencies of features, up until this dissertation, MDA had only been applied to text samples above at least 120 words (e.g. Sardinha, 2014), meaning that MDA research has been limited into investigating longer texts, or short texts, which have been concatenated to a size more suitable for frequency-based analyses.

Additionally, exploratory factor analysis has been the most common statistical method used for this type of analysis; however, other statistical methods have been used, including confirmatory factor analysis (Biber, 2007), principal component analysis (Passonneau et al., 2014), and discriminant analysis (Sardinha and Pinto, 2015; Biber and Egbert, 2016), all of which are appropriate for frequency-based analyses.

This dissertation is the first to offer a solution to analysing the major patterns of linguistic co-occurrence in short texts, like tweets. The solution involves subjecting the occurrence of linguistic features (i.e. presence or absence), as opposed to their relative frequency, to MCA instead of factor analysis to identify the major patterns of linguistic co-occurrence. MCA analyses categorical variables, as opposed to continuous variables like the relative frequencies of features. When MCA is applied to the presence/absence of linguistic features in tweets, it identifies the tweets that are similar to each other with respect to their linguistic form, and it identifies relationships between the linguistic features; that is, those that co-occur together and those that rarely co-occur in the tweets. In this way, MCA is used analogously to factor analysis with respect to its role in MDA.

The results of the MCA-for-MDA presented in this dissertation have been coherent and interpretable for both general Twitter and Twitter trolling, suggesting that it is a valuable alternative technique to factor analysis for the MDA of short texts. MCA requires fewer decisions before the interpretation than factor analysis. For example, the amount of factors extracted and interpreted in factor analysis need to be decided on before they are extracted and changing this amount could influence the features that load on them. The dimensions returned in MCA, however, are stable and do not change if you extract 3, 4, 5 or 80. Additionally, unlike factor analysis, particular rotation methods need not be applied in MCA to observe the strongest patterns (Biber, 1988). Nevertheless, it is important that one interprets a dimension by looking at each feature across the dimensions, as despite contributing strongly to a dimension, one feature may contribute considerably more to another dimension. This is not to say that factor analysis should not be used. Factor analysis is extremely powerful for investigating language variation across long texts. Instead, it is argued here that MCA is a great alternative for the analysis of short texts, and that the dimensions are more stable.

MCA can return as many dimensions as there are categories of linguistic features minus the amount of linguistic features. For example, the MCA-for-MDA of general Twitter returned 63 dimensions because 63 linguistic features were analysed each with two categories (presence/absence). Due to the high dimensionality, the percentage of variance explained by the dimensions tends to be lower than other dimension reduction methods, although there is a formula to retrieve a better percentage for the first few dimensions (see section 4.7). This transition formula was applied to both sets of results. As a result, the percentage of variances explained by the first and second dimensions of both studies were inflated and the other dimensions were revealed to explain less amount of the variance. Whilst this essentially confirms that tweet length is the strongest influence on the presence or absence of features, it does not elucidate how much of the data is explained by the other dimensions.

In addition to enabling a short text version of full MDA, MCA has also facilitated a short text version of projecting texts onto existing dimensions through the use of supplementary elements, which has enabled a systematic comparison of the distribution of trolling tweets in relation to the dimensions of general Twitter. This broadens the scope of MCA-for-MDA for comparing different kinds of short texts to other shorts texts, like the comparison of other kinds of tweets to general Twitter's dimension.

The main problem is that text length is not controlled for in this modified approach because the relative frequencies of features are not taken. Nevertheless, it is possible to observe the degree to which the dimension patterns are correlated to the length of the texts. Fortunately, for both MDAs it was only correlated strongly to the first dimension with a slight positive correlation to the second dimension of both studies, showing that informationally dense broadcasting tweets tend to be longer than interactive tweets. Overall, this thesis has demonstrated the power of MCA, not only for achieving short text versions of both kinds of MDA, but more generally for the purpose of reducing a large categorical data set to a smaller number of underlying dimensions. Consequently, MCA has particular uses across linguistics and phonetics, especially when the variables are categorical in nature.

Although this research offers a solution to applying MDA to short texts, tweets are consistently short, in that they rarely exceed 40 words. Future research may seek to apply this to other forms of social media or other platforms where short texts occur. Whilst short texts occur on social media and other platforms, they are not always short nor are they restricted to being short in the same way as tweets. For example, Reddit users can post messages that are both exceptionally large and exceptionally small. As posts get longer, the analysis of the occurrence of features may not be as informative, as the likelihood of observing a linguistic feature increases with the more words, as was observed in Dimension 1. Consequently, future research should invest in developing ways to analyse texts from platforms that are of varying lengths, including exceptionally short and exceptionally long in one study.

8.3.2. The Tagger

To investigate the major patterns of linguistic co-occurrence across the texts of a corpus, it is necessary that the linguistic features analysed represent the language variety, and that the tagger is accurately identifying these features. Biber (2019: 14) emphasises the importance of the computer program used to grammatically tag the texts of the corpus, as this “provides the foundation for MD studies”. The tagger that Biber has developed has been revised over the years, incorporating new and more detailed features, and it has also been extended for various language varieties by incorporating features specific to the language variety under investigation. Unfortunately, Biber’s grammatical tagger is not well suited to tweets, which tend to include lots of non-standard grammar and spelling. Moreover, the feature set of this tagger is also not tailored for features specific to CMC and Twitter. There are features, which can occur in tweets, such as hashtags, mentioning through the ‘@’ symbol, pronoun omissions and URLs, which are functionally important, and therefore these need to be investigated. Given that this thesis sought to conduct an MDA of tweets (trolling and general English tweets), a new tagger was created, which not only accounted for spelling variations and non-standard grammar, but also incorporated features unique to this language variety.

This thesis offers a computer program for tagging tweets for the MDA feature set and features related to tweets and CMC, which has been evaluated for its accuracy. Whilst it is not 100 percent accurate on every feature, it achieves levels of accuracy that are on par with other taggers with a similarly detailed feature set. Of course, it would be ideal to achieve perfect accuracy, but the reality is that there will be errors and overall these errors are expected, even more so with data like Twitter. Gray (2019) notes that it is up to the researcher to investigate how accurate the tagger is and what level of accuracy is acceptable. For instance, this dissertation was counting the presence/absence of tags, and so the accuracy of the tagger mainly concerned the precision of the tag, as opposed to the tagger's recall rate (i.e. missed features).

This tagger is a unique contribution to other MDA studies that have looked at tweets (Titak and Roberson, 2013; Friginal, Waugh and Titak, 2018; Sardinha, 2014, 2018). These studies have offered limited explanation about the noisiness of tweets and other forms of social media/CMC (e.g. Facebook), how the noisiness impacts the accuracy of the tagger, and what steps are taken to reduce the effect. Moreover, there has been little attempt at representing this language variety in the feature set. For instance, Titak and Roberson (2013), Friginal, Waugh and Titak (2018) and Sardinha (2018) have not incorporated features like '@' mentioning, URLs, and emojis in their investigations of web registers, where these features have been found to frequently occur. The most important critique of MDA studies was offered by McEnery et al., (2006), which emphasised that the feature set needs to be all-inclusive and representative. This thesis has aimed to incorporate these specific linguistic features found across Twitter. Nevertheless, there are limitations of the tagger used in this thesis. First, grammatical taggers reflect the analyser/creator's interpretation or theory of grammar (McEnery et al., 2006). For the most part, the tagger is based on the particular grammar utilised across most MDA studies, which is the Longman Grammar of Spoken and Written English (e.g. Biber et al. 1999). However, there are additional features, especially as a result of pooling (see Appendix 2), which are potentially reflective of my theory of grammar. For reasons of transparency, the tagged corpora and the tagger are available upon request, meaning that the grammar is explicit and recoverable to anyone that is interested.

Another limitation is that it does not incorporate some of the features that are found in the feature sets of more recent MDA studies (e.g. Sardinha, 2018), such as various noun types like abstract nouns, technical/concrete nouns, and cognitive nouns. Additionally, more recent studies have distinguished more specific kinds of features. For example, the tagger developed for this thesis distinguished *that* and *to* complement clauses according to the part-of-speech tag that it was controlled by (e.g. adjective, noun, verb). However, more recent studies take this further and distinguish the different kinds of *that* and *to* complement clauses based on the specific verbs, adjectives or nouns by which they are controlled. For instance, Sardinha (2018) distinguishes *that* complement clauses that are controlled by adjectives denoting likelihood from those controlled by factive adjectives, among many others. Fortunately, features so specific like the different kinds of complement clauses used in Sardinha (2018) described above are too rare across the Twitter corpus used in this dissertation. As mentioned, this thesis only distinguished complement clauses according to the part of speech tag that they were controlled by (e.g. adjective *to* complement clauses), as opposed to specific types of the part of speech tag (e.g. factive adjectives), and even then these individually were still too rare to be included as single features in the final feature set. Given the rarity of complement clauses on Twitter distinguished by the broad part of speech that controls it, it can be argued that more specific types would be just as rare. Thus, whilst the tagger is not as detailed with respect to some features as more recent MDA studies, it was not always required given the rarity of many features on Twitter. Nevertheless, the tagger is limited and future research should attempt to be as inclusive as possible and allow the data to tell if the features are rare.

8.3.3. Dealing with Infrequent Features

In addition to the tagger, this dissertation offers a transparent feature set and a clear approach for dealing with infrequent features and the decisions and justifications made for this analysis. Appendix 2 reveals the steps taken and decisions made when features were rare. When it was possible, rare features were systematically pooled with other similar features into a broader grammatical category. The effects of pooling were tested by running separate short text MDAs on the tweets analysed for the occurrence of the rare feature pooled to a broader category or with the rare feature removed from the feature set altogether. These results were correlated to each other to observe whether it led to any dramatic effect.

Rare features needed to be treated this way, given the way that MCA deals with rare features (Le Roux and Rouanet, 2010). Pooling may be perceived negatively in the sense that it loses the distinction of particular features, especially considering that these distinctions are made in the first place because they are important for particular functions. For example, in the general Twitter study agentless passives occurred in more than 5% of the tweets, whereas by-passives did not. One approach would be to maintain the distinction and remove by-passives from the feature set and only investigate how agentless passives co-occur with other features. This approach, however, ignores cases of by-passives. Ignoring structures can mean that certain tweets will not be fairly represented in the dimensions. For instance, consider a tweet that says: “Nigel Farage was attacked by Paul Crowther”. If the analysis ignores by-passives, then this tweet contains only proper nouns and past tense verbs. Whilst it may be associated to a particular dimension, other tweets may be more associated to the dimension, especially ones containing more of the features comprised in the linguistic co-occurrence patterns in the dimension. Additionally, removing features because they are rare implies that each case of the feature is functioning differently. With respect to passive constructions, by-passives and agentless passives can have different functions. Nevertheless, they also both have an over-arching function as passives. Thus, another approach is to pool features, aggregating similar features into a broader tag. Whilst pooling loses the distinction, one is able to make higher-level interpretations when features are pooled, as the more general communicative function can be assumed. Moreover, if pooling occurs and no difference is found with respect to the co-occurrence patterns, and tweets are similar in their associations to the patterns, it suggests that the broader function of the feature is more likely being realised in the context of the tweets because if each type of feature was providing different functions, then they may pattern with different features and thus the pattern of the dimension would not be consistent.

The method used in this thesis provides an easy way to assess the effect of pooling. For the most part, the results suggest that pooling has little effect on the contributions/coordinates of the features and the coordinates of the texts, which indicates that the functions of the features are broader, as was also observed in the realisations of the features in context. Overall, pooling has enabled higher-level interpretations.

8.3.4. The Method for Assessing Representativeness

Given previous studies examining tweets, one may argue that the corpora of tweets are small and therefore non-representative. The assumption that large corpora are representative is misguided and largely depends on one's research goals and definition of representativeness. Egbert (2019) describes a survey of Corpus Linguistic studies that he conducted with Bethany Gray and Douglas Biber to explore common conceptualisations of *representativeness* and *corpora.* Additionally, they investigated the degree of detail offered in the description of the corpus and the ways that researchers evaluate representativeness in order to assess the degree to which Biber's (1993) suggestions have been integrated into the field. Specifically, Biber emphasises that "representativeness refers to the extent to which a sample includes the full range of variability in a population" (1993: 244). Overall, Egbert (2019) describes how their survey results show that this has not been integrated, where not only did they find studies with limited details with respect to defining corpora and representativeness, but they also found that representativeness was not a primary concern of the researchers. However, in studies where it was a concern, target domain representativeness was focused on, whereas linguistic distributions were not. Egbert (2019) notes that very few corpus linguistic studies in their survey evaluated the representativeness of the corpus, and instead they tended to emphasise the importance of size.

As a result of these shortcomings, Egbert (2019) proposes a nine-step process for designing and collecting a representative corpus and describes two case studies. These case studies clearly illustrate many of the steps; however, not all of them are represented, especially the most important, which is the process for evaluating linguistic representativeness. Egbert (2019) instead argues that they did not need to evaluate their representativeness because the studies were examining high frequency features. Biber (1993) suggests that studies examining specialised discourse domains and high frequency items do not need to be large as these tend to be stable in smaller samples (i.e. less internal variation requires fewer texts). It is in my view that this should not necessarily be assumed and that this should be tested and evaluated.

Few studies have attempted to evaluate linguistic representativeness, and few studies have offered methods for doing so. This dissertation may therefore be taken as a response to Egbert's (2019) step-by-step process and also as a step towards developing new ways to assess and conceptualise linguistic representativeness. This dissertation offers a method for assessing the representativeness of linguistic distributions by conceptualising it as dimension stability in smaller samples of the larger corpus. In particular, when smaller samples of the larger corpus are correlated according to the patterns of linguistic variation, then it can be argued that the linguistic distributions are stable and that there are enough texts, which represent the range of linguistic distributions. When smaller samples are not correlated, it indicates that the linguistic distributions are not stable and that more texts are needed in order to observe the full range of linguistic distributions. This dissertation found that the major linguistic distributions found in the two corpora of tweets were stable in samples that were less than half of the overall corpus. The linguistic distributions referred to here were concerned with grammatical variation, and thus it is important to stress that these corpora of tweets are not representative of the words that are found across English Twitter and Twitter trolling. To have a corpus representative of the words found across Twitter, more texts would need to be collected. Overall, this approach provides a method for assessing dimension stability and representativeness, defined by the range of linguistic distributions in the population.

8.4. Limitations

8.4.1. The Data

Although it is argued that the general Twitter and Twitter trolling corpus are representative of the range of linguistic distributions across them, both corpora are limited. The general Twitter corpus is limited because it only includes four hours of English tweets collected on one day from the random 1% stream of public tweets. A corpus of tweets spanning a longer time frame may reveal different, but most likely more dimensions of linguistic variation. Future research should therefore seek to test whether these dimensions of linguistic variation are also observed in a corpus of tweets collected over a long time period. Moreover, it would be interesting to observe whether the language of Twitter changes over time in a similar way to Clarke and Grieve (2019) but across general Twitter.

The trolling corpus is limited because it is based on what other people have perceived. Given that trolling is largely a deceptive practice, some types of trolling may be more likely to be accused of trolling, whilst others may not. For example, based on common media descriptions of trolling that emphasise its hostile and antagonistic nature, the highly uncivil kinds of posts may be perceived as trolling more often than the civil kinds. As a result, the data may be biased towards particular kinds of trolling. Additionally, sarcastic accusations are always a possibility, but given Poe’s Law (2005), it is impossible to tell. It is also not clear if this corpus contains self-identifying trolls, like Phillips’ (2016) research. Whilst it would be preferable to have a corpus where each trolling post was definitely known to be trolling, the reality is that the Internet is full of mischief and identity-play, meaning that no one will ever know who is being serious and who is not. This was the least biased way of collecting trolling posts, but it is nonetheless limited.

Overall, the dissertation is also limited in that it is only focused on trolling on Twitter and not across the entire Internet. Like Twitter, the language across different platforms and websites will be constrained and influenced by the technological affordances, the users and the community norms, and by extension, these will also likely influence the trolling on those platforms and their major linguistic repertoires. Future research should attempt to run similar analyses on trolling across other web-based platforms to observe if these linguistic repertoires are generalisable across them, and also to observe if trolling on the platforms also exploit the major dimensions of linguistic variation of general social media posts.

8.4.2. The Feature Set

The analysis finds that trolling tweets and general tweets are largely indistinguishable according to General Twitter’s dimensions of linguistic variation. However, these dimensions are based on the feature set used and this feature set is not all-inclusive of every single linguistic feature. This suggests that trolling tweets may be more distinguishable from general tweets with different and more features.

For example, semantic information is only partially incorporated, such as the inclusion of private and public verbs. Semantic information could bring an additional dimension to the results. For instance, consider the semantic category of ‘DEATH and DESTRUCTION VERBS’, which includes verbs like murder, kill, die, obliterate, etc. If such a semantic category was incorporated into the feature set of MDA, it may enable the distinction between more aggressive and antagonistic communicative functions with more positive communicative behaviour. The results show that Twitter trolls exploit the major patterns of functional linguistic variation for subversive ends. Semantic information may help in the process of identifying the particular ways in which they do this.

8.4.3. Construct Validity

One of the most difficult steps in MDA is assigning an interpretative label that provides the best description of the linguistic co-occurrence patterns represented in the dimensions of functional linguistic variation. It is important to note that the interpretations do not influence the results of the MCA; rather the labels are constructs assigned. It is probable that one may disagree with the labels assigned to the dimensions in this dissertation. However, these dimensions have been rigorously interpreted and the best labels in my view have been assigned, which explains the patterns of linguistic co-occurrence on both sides of the dimensions.

Despite this, the labels have not been validated. Pavalanathan et al. (2017) proposed different ways to validate the construct assigned to their dimensions of frequently co-occurring stance markers both intrinsically and extrinsically. One intrinsic approach they used was a word intrusion task, which encouraged coders to identify a word that was intruding based on the notion that a coder will be able to identify an intruder when the target concept is internally consistent. This dissertation is not dealing with words per say and thus this approach is not appropriate. One extrinsic approach involved using previously annotated corpora of the particular interpretations to see if such stance markers also occurred in these (e.g. polite and formal corpora). Future research could set up other experiments to validate the labels assigned to the dimensions of linguistic variation described here. For instance, one could get annotators to group texts known to be associated to the particular linguistic co-occurrence patterns under the relevant label. Another way may involve other researchers simultaneously interpreting the dimensions and comparing the reliability of the labels assigned. Despite not implementing any of these approaches for validating the labels assigned in this dissertation, this research has been presented at various conferences and the interpretations have not yet been disputed.

9. Conclusion

This dissertation provides a quantitative linguistic description of Twitter trolling, not only in terms of its major dimensions of linguistic variation, but also with respect to general Twitter. The major dimensions of linguistic variation of the Twitter trolling corpus provide empirical evidence that trolling tweets vary according to how interactive or informational they are. Trolling tweets also vary according to how personal their focus is or if they are externally focused. Additionally, trolling tweets vary according to their degree of a promotional or oppositional style, and with respect to their level of civility or incivility. As the major dimensions of linguistic variation of this corpus of trolling, it suggests that these major linguistic repertoires are important communicative styles for trolling on Twitter. Apart from Dimension 5 (the degree of civility), the results show that the major dimensions of linguistic variation are shared across both corpora, despite them being collected with different constraints. In addition to sharing many of the major dimensions of linguistic variation, the results of the comparison of trolling tweets in relation to general Twitter's dimensions show that trolling tweets vary in much the same way as general tweets along all of General Twitter's dimensions of linguistic variation, which shows that, overall, trolling tweets are remarkably more similar to general tweets than they are different.

In addition to partially informing theories that trolling is a deceptive practice (Donath, 1999; Cruz et al. 2018), the substantial similarity between trolling tweets and general tweets may also be partially explained by the idea that some trolls are détourning the spectacle (Phillips 2013; Debord, 1967). Détournement derives from Debord's (1967) book "The Society of the Spectacle", which is a critique of modern day consumer culture, commodity fetishism, an obsession with images, and a lack of authenticity. In this book, Debord (1967) introduces the spectacle, which is a reflection of society, where relations between people have been taken over by relations to commodities. In particular, Debord (1967) critiques modern consumer culture and commodity fetishism, as humans have degraded from being into having because the spectacle communicates what people must have through images and spectacular language. As a result of consumerism and the spectacle, social life has moved from having into a state of appearing and showing what you have, especially through images and spectacular language. According to Debord (1967) the spectacle is at the heart of inauthenticity of an authentic and real society (1967: 6). The spectacle is merely a representation of social life. "The spectacle presents itself as something enormously positive, indisputable and inaccessible. It says nothing more than "that which appears is good, that which is good appears" (Debord, 1967: 12), and it demands only passive acceptance, which, according to Debord, is already obtained because just by appearing it is accepted as true.

According to Debord (1967) lived reality is invaded by the thought of the spectacle. Consequently, to restore authenticity and shine the light on the spectacle, Debord (1967) encouraged the use of détournement. Détournement involves using the same kind of language and images associated with the spectacle in order to disrupt and mock the spectacle. It is also known as culture jamming, where particular behaviours from popular culture are recontextualised or appropriated and given subversive and antithetical meanings (Phillips, 2013). One example of culture jamming occurred in 1977 when the Sex Pistols played their version of “God Save the Queen”, whilst on a boat sailing along the River Thames in order to mock the Silver Jubilee river procession that was in honour of Queen Elizabeth II (Spencer, 2012).

This dissertation shows that trolls are drawing on the same major linguistic repertoires as general tweets for trollish and subversive ends. For example, the analysis showed that some trolls draw on an informational broadcasting style to spread hyperbolic misinformation. Additionally, some trolls draw on a supportive and promotional communicative style to express support for transgressive ideologies. Moreover, the results show that some trolls employ an interactive style in order to mock their addressee. Thus, in light of Debord (1967) I argue that some trolls are détourning the spectacle, which, based on the patterns of linguistic variation, is the use of social media for self-commodification.

The use of détournement to describe trolling practices is not a new idea. Phillips (2013) argues that the media has a huge amount in common with trolling. Although the media’s and trolling’s ends are very different, Phillips (2013) finds that the means for achieving those ends are near enough identical, mainly because these means are about gaining attention and being successful. On the one hand, the media must gain the attention of the public so that they can make a profit, as attention equates to advertising revenue (Phillips, 2013). On the other hand, trolls must gain the attention of their victims so that a response can be obtained for the accrual of lulz (Phillips, 2013). To attract attention, Phillips (2013) argues that trolls and the media engage in spectacle. Spectacle in Phillips' (2013) view is sensationalistic and exaggerated content, that can often be so detached from truth, that it becomes only an appearance or representation of what actually happened. Trolls jam the culture of the media by embodying their sensationalistic spectacular nature to the point where the trolls behaviour plays right into the media's hands, and the media are compelled to report on the trolling behaviour because it fits perfectly with the demand for sensationalistic content, which attracts attention and keeps advertising revenue high (Phillips, 2013). Phillips (2013) suggests that this causes trolls to howl with laughter because the media are often critiquing the trolls for their behaviour, but because the trolls are détourning media practices, in actual fact, the media are unknowingly critiquing themselves. The purpose is to show that the media's behaviours tend to be accepted as essential for profit-making, and yet trolling behaviours, which are exactly the same, are condemned (Phillips, 2016). By détourning the spectacle of the media, trolls essentially point the finger at mainstream media culture and their inauthentic representation of events to get a profit (Phillips, 2013).

The same might be said for some instances of trolling on Twitter and social media more generally. This dissertation finds that Twitter trolls are employing many of the same linguistic repertoires as general tweets in order to mock Twitter culture and for a variety of subversive ends. By détourning the major linguistic repertoires, trolls are pointing the finger at social media culture – at least by proxy (Phillips, 2013). Echoing Debord (1967), I argue that the spectacle of social media is an inauthentic representation of social life. In the shift from Web 1.0 to Web 2.0, social media sites grew at a rapid pace and invaded social life, affording the opportunity to be active, generating our own content, as opposed to passively consuming content like news (Seargeant and Tagg, 2014). As people created content, they mapped their identities onto it, promoting themselves, their brand, their opinions and their everyday lives. But online identities do not live, they only appear for the consumption of others. And they are only consumed by the people who pay attention, leading content creators to engage in spectacle – aimed at gaining the attention of others.

Social media posts supposedly represent their author's lives, but for the most part they report on good things, or things that present themselves in a positive light - their achievements and their successes – and the things that gain them attention. Their posts are only mere representations of them and what is going on in the world. Twitter and social media platforms more generally have become the shop window, where the products are our posts and by extension our identities, curated for the consumption of others in order to gain some form of capital (economic or non-economic).

The spectacle may therefore be the appearance of self - the presentation of one's best, most interesting and idealistic self. The self merely appears on social media using spectacular language and images to create captivating identities for the consumption by particular audiences (Baym, 2014), whether that is one's followers, one's addressee(s), or whether it is a far broader intended or imagined audience (Schmidt, 2014). If some cases of trolling are détourning the spectacle, trolling may be positioned as subversive and an attack to the self-commodification and inauthenticity on social media and web 2.0. sites. Trolls use spectacular language to similarly present transgressive ideologies and misinformation, and create transgressive, deceptive and inauthentic identities for the consumption of their victims in order to gain their attention and provoke a response. In a context that is so tied to one's carefully crafted identity and "best self", trolling, which is often aimed at humiliating and triggering their victims into responding (often in a negative way), becomes so much more troublesome and arguably so much more effective, as it is an attack on one's idealistic commodified self. Moreover, trolling is effective because it often brings out their victim's worst self, as their victim's response can be negative, as they denounce, correct, counter, and critique the troll. This suggests that vanity can be conducive to trolling, as some trolls seek to humiliate anyone, especially those who fathom themselves important because the effect is stronger.

Social media users and these kinds of trolls, then, are both invested in the spectacle (Phillips, 2013). Social media users are invested in the accrual of capital (economic and social - likes, retweets, followers), whilst trolls are often invested in the accrual of lulz (Phillips, 2013). To accrue capital, social media users engage in spectacle - practices of self-branding, broadcasting, interacting publicly, self-reporting, promotion, persuasion, and stance taking, all of which contribute to the creation of an appearance of the self - an inauthentic appearance of self that gains attention. To accrue lulz, some trolls embody these communicative practices and behaviours, but exploit them for their trollish and subversive needs, so that when they are critiqued they can point the finger at general social media users. Thus, if some trolling is hijacking the spectacle of social media and turning it back on itself, perhaps then, trolling in this case is not the problem. Perhaps trolling is the symptom of the spectacle. Perhaps trolling, like symptoms of a disease, can enable the diagnosis of the problem, which may be the commodification of an inauthentic representation of social life on social media.

Notably, this dissertation began by emphasising the diversity of trolling. Trolling has come to mean different things for different people in different contexts and as a result it has become a term that encapsulates a whole range of behaviours and purposes. This dissertation collected trolling posts using other people's perceptions to better represent this diversity. The analysis demonstrates this diversity, revealing major linguistic repertoires and showing that trolling is so much more than provoking a response and mocking their victims. Whilst détournement and deception are able to explain part of the overlap between the major patterns of linguistic variation of trolling tweets and general tweets, these theories tend to account for stereotypical kinds of trolling, as opposed to all kinds of trolling. Importantly, not all instances of trolling are trying to provoke a response or mock their victims. Thus, not all instances of trolling are purposefully tweeting in the same way as general tweets in order to mock general Twitter users (i.e. détournement) and not all trolls are trying to deceive their victims into thinking they are genuine in order to provoke a response (i.e. assimilation). Trolls may have different goals and troll for a variety of purposes, using language in particular ways in order to pursue these interests. For example, many Twitter trolling posts in the corpus are, among others, insulting, silencing, attacking, opposing and misleading their victims. The findings of this dissertation reveal the major communicative functions of trolling tweets and show how these can be exploited for these diverse trollish ends. Thus, to conclude there is not a one-size-fits-all explanation for the overlap between trolling tweets and general tweets. Overall, it is argued that trolls draw on the major communicative functions and patterns of linguistic variation of general Twitter but exploit them in order to pursue their individual interests.

Appendices

Appendix 1: The Feature Sets and Decisions for Pooling

Le Roux and Rouanet (2010) advise that very infrequent features (e.g. those that occur in < 5% of the data) either need to be pooled with other related features or they might need to be discarded because infrequent features can overly influence the axes, as they contribute more to the overall variance. Table 49 presents the features occurring in fewer than five percent of the tweets in the general English Twitter corpus and the decision that was made with respect to pooling or deleting features from the final dataset. The justifications for these decisions are also presented in the Table. Infrequent features that were specific types of a broader part-of-speech category were pooled into the broader part-of-speech or 'other' category. For example, the feature set distinguishes between different kinds of adverbs (e.g. place, time, downtoner, amplifier, *see Table 3*), and then any other adverbs that are not one of these types are tagged as 'Other Adverb'. Quantifying adverbs do not occur in more than five percent of the tweets. Therefore, this feature was pooled with the 'Other Adverb' category. Essentially, because it does not occur frequently, the feature is dropped from the feature set, meaning that if quantifying adverb wasn't included in the tagger, any occurrence of a quantifying adverb would be classed as an instance of 'Other adverb'. Thus, this is the logical category with which to pool it. When there was more than one option (deleting or pooling or pooling the feature into multiple categories), many of these options were tested by running several MCAs on different feature sets depicting the different pooling options. For example, copular verbs that are not BE as a main verb did not occur in more than five percent of the tweets, whereas BE as a main verb did. Both features are part of the broader category of ‘stative forms’, and so they could be pooled together into one broad category, or copular verbs could be deleted from the feature set. To test the effect of either decision, two data matrices were created and each was subjected to MCA: one with all the other linguistic features but copular verbs were deleted, and the other involved copular verbs being pooled with BE as main verb into the new category of ‘stative forms’. Although the active variables in each MCA are different, the individual tweets are the same, meaning that they can be compared. Consequently, the coordinates and contributions of the individuals in each MCA were correlated to the other to observe if there was a substantial difference between the two feature sets. For the most part, the decision to delete a feature or pool it with other categories or broader features made little difference to the position of tweets, where the dimensions (at least the first 10) from one MCA were strongly positively correlated to the corresponding dimensions in the other MCA in regards to the contributions ($r > .85$) and coordinates ($r > .91$) of the individual tweets. The one pooling feature that made the most difference was by grouping all kinds of initial verbs into one category of ‘initial verb’, which resulted in weak and moderate correlations between the corresponding dimension 3 and dimension 4 coordinates ($r < .20$) and contributions ($r < .55$).

This substantial difference might have meant that this feature needed to be included. Therefore, a full interpretation of these dimensions in both feature sets (including or excluding initial verbs) was conducted. The dimensions of the feature set that included initial verbs, especially those dimensions where initial verbs was a strongly contributing feature, were not as easily interpretable as the dimensions returned from the MCA on the feature set excluding all initial verbs. This is arguably because initial verbs have many functions and some instances of initial verbs (e.g. initial verbs ending in -ing) may not even be verbs, but may be gerunds. In addition to issues of interpretability, the percentage of variance explained by the major dimensions on the feature set including initial verbs was less than the percentage of variance explained by the major dimensions from the feature set excluding this feature. For these reasons, 'initial verb' as a broad category was excluded from the final feature set.

Table 49: The feature occurring in fewer than 5% of the general English tweets, and the decisions and justifications for inclusion/exclusion in the final feature set

Features < 5 % of General English Tweets	Decision	Justification
Acronym	Deleted	Acronyms are already counted as the part of speech tag assigned (e.g. U.S.A_^ is assigned proper noun tag)
Adj+that complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.

Adj+to complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.
Adverbs of frequency/ usuality	Pooled with General Adverbs	Adverbs are divided into different types and all other adverbs not specified are grouped into a broader 'other adverbs' category. If the specific type of adverb occurs infrequently then each instance can be re-combined with the 'other adverb' feature.
Agentless Passives	Pooled with Agentless Passives into passive constructions	Passive constructions were divided into different types and therefore re-combined to form the broader category.
Bracket	Deleted	No applicable broader category, albeit 'punctuation'; however this would lead to the loss of the distinction between question marks and exclamation marks.
By-Passive	Pooled with By-Passives into passive constructions	Passive constructions were divided into different types and therefore re-combined to form the broader category.
Cause Subordinator	Pooled with General Subordinator	Subordinators are divided into different types and all other subordinators are grouped into an 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.
Comparative	Pooled with Superlative in Gradation	I combined with comparatives to form a broader 'gradation' category, as both comparatives and superlatives occur too infrequently to stand on their own but occur enough when combined. Both forms can be used to form an evaluation.
Concessive Subordinator	Pooled with General Subordinator	Subordinators are divided into different types and all other subordinators are grouped into an 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.
Conditional Subordinator	Pooled with General Subordinator	Subordinators are divided into different types and all other subordinators are grouped into an 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.

Copular verbs (not BE)	Pooled with BE as main verb	I could have deleted this feature from the feature set, or combined either with ‘General/Other Verbs’ or pooled with BE as a main verb and existential there into a category of Stative forms. I tested all by running the analysis on all types and then correlated the coordinates of the individual tweets from the results. I also compared the contribution and coordinate of ‘Stative forms’ and ‘BE as a main verb’, as well as ‘General/Other verbs’ pre- and post- pooling in each analysis on each dimension to observe if the pooling led to any substantial difference.
Downtoner Adverb	Pooled with General Adverbs	Adverbs are divided into different types and all other adverbs not specified are grouped into a broader 'other adverbs' category. If the specific type of adverb occurs infrequently then each instance can be re-combined with the 'other adverb' feature.
Existential *there*	Pooled with BE as a main verb and Copular verbs	I could have deleted or combined with BE as a main verb and copular verbs into a category of 'stative forms'. I tested both by running the analysis on both scenarios and then correlated the coordinates of the individual tweets from the results on all the dimensions. I also compared the contribution and coordinate of ‘Stative forms’ and ‘BE as a main verb’, in each analysis on each dimension to observe if the pooling led to any substantial difference.
Gerund	Deleted	No applicable broader category.
Initial verb	Deleted	Whatever the verb is, it would also be classified as either one of the verb types or in the 'other verb' category. Therefore it does not need to be pooled with broader verb category. I could have combined with other initial verbs. However, I tested this by running the analysis on the feature combined with other initial verbs as well as with this feature deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb -S	Deleted	This feature is already counted as third person singular verb form regardless of initial position. I could have combined it with other initial verbs. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.

Initial verb BE	Deleted	I could have combined into a broader category of initial verbs with other initial verb instances. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb DO	Deleted	I could have combined into a broader category of initial verbs with other initial verb instances. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb HAVE	Deleted	I could have combined into a broader category of initial verbs with other initial verb instances. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb ING	Deleted	It could be an auxiliary omission and thus in progressive form or it could be a gerund. Rather than check each instance to clarify and rather than group these instances into either general verbs or general nouns and misclassify some, it was decided to just delete the feature altogether from the feature set. I could have combined with initial verbs (but it might not have been one). However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb MODAL	Deleted	The type of modal is counted any way regardless of whether it is positioned initially. I could have combined it with other initial verbs. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.

Initial verb PAST	Deleted	This feature is already counted as past tense verb regardless of initial position. I could have combined it with other initial verbs. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb QUESTION	Deleted	Too few instances to combine with other question features. I could have combined with other initial verb types. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Laughter Interjection	Pooled with General Interjections	All specific types of interjections occurred in fewer than five percent of tweets. These specific types were therefore combined into a broader category of interjections.
Modal of Necessity	Deleted	I could have combined all modals together but this would lose the distinction between possibility and prediction modals, which contribute differently on different dimensions.
Negative Interjection	Pooled with General Interjections	All specific types of interjections occurred in fewer than five percent of tweets. These specific types were therefore combined into a broader category of interjections.
Noun+that complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.
Ordinal determiner	Deleted	Not enough instances of either kind of ordinal to combine to create an ordinal feature. I could have combined with other determiners (e.g. definite and indefinite article and demonstratives and possessive) but this would lose the distinction between them.
Ordinal Noun	Deleted	Not enough instances of either kind of ordinal to combine to create an ordinal feature. I could have combined to general noun category, but because I deleted ordinal determiners from the feature set I also decided to delete ordinal nouns.

Other Conjunctions	Deleted	Too few instances to combine with other frequent and more specific kinds of conjunctions.
Pied-piping relative	Deleted	Not enough instances of either kind of relative clause to combine into a broader category of relatives
Place Adverb	Pooled with General Adverbs	Adverbs are divided into different types and all other adverbs not specified are grouped into a broader 'other adverbs' category. If the specific type of adverb occurs infrequently then each instance can be re-combined with the 'other adverb' feature.
Place subordinator	Pooled with General Subordinator	Subordinators are divided into different types and all other subordinators are grouped into an 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.
Positive Interjection	Pooled with General Interjections	All specific types of interjections occurred in fewer than five percent of tweets. These specific types were therefore combined into a broader category of interjections.
Possessive Noun	Pooled with Possession	The only feature denoting possession that occurred in more than 5% of the tweets was possessive determiner. These features were combined to create one variable of possession.
Possessive Pronoun	Pooled with Possession	The only feature denoting possession that occurred in more than 5% of the tweets was possessive determiner. These features were combined to create one variable of possession.The type of possessive pronoun is also counted (e.g. first, second, third or IT).
Possessive Proper Noun	Pooled with Possession	The only feature denoting possession that occurred in more than 5% of the tweets was possessive determiner. These features were combined to create one variable of possession.
Pre-Determiner	Deleted	Too few instances to combine with quantifying pre-determiners to create general pre-determiner category.
Pro-verb DO	Pooled with General verbs	Different types of verbs were distinguished from general verbs and therefore infrequent types can be recombined with broader verb category.

Quantifying Adverb	Pooled with General Adverbs	Adverbs were divided into different types and all other adverbs that do not fall in these particular categories are grouped into a category called 'other adverbs'. Therefore, if a specific type does not occur frequently it can rejoin the 'other adverbs' category. Quantifying adverbs could have also been grouped with other quantifiers of different parts-of-speech (e.g. quantifying -determiners, -pre-determiners, -pronouns). They were not grouped this way because quantifying determiners and pronouns occurred in more than five percent of general English tweets and I did not want to lose this part-of-speech distinction by grouping them all together.
Quantifying Pre-Determiner	Deleted	Too few instances to combine with other pre-determiners to create general pre-determiner category.
Reflexive Pronoun	Deleted	No applicable broader category, albeit 'pronouns'. Reflexive pronouns are counted according to first, second or third person, or IT.
Relative Clause Object Gap	Deleted	Not enough instances of either kind of relative clause to combine into a broader category of relatives
Relative Clause Subject Gap	Deleted	Not enough instances of either kind of relative clause to combine into a broader category of relatives
Semi-Colon	Deleted	No applicable broader category, albeit 'punctuation'; however this would lead to the loss of the distinction between question marks and exclamation marks.
Split Infinitive	Pooled with infinitives	Split infinitives were separated from infinitives as a particular type and so therefore were recombined to the broader category.
Suasive verb	Pooled with General verbs	Different types of verbs were distinguished from general verbs and therefore infrequent types can be recombined with broader verb category.
Subordinator with ellipted subject	Deleted	No applicable broader category. If it is a specific type of subordinator it will be classified as such as well.
Superlative	Pooled with Superlative in Gradation	I combined with comparatives to form a broader 'gradation' category, as both comparatives and superlatives occur too infrequently to stand on their own but occur enough when combined.

Synthetic negation	Deleted	No applicable broader category, albeit 'negation', meaning that I could have combined analytic negation with synthetic negation. However, I did not want to conflate this distinction as previous research has found this to be an important feature (e.g. Biber, 1988; Clarke, 2018; Clarke and Grieve, 2017).
Time Subordinator	Pooled with General Subordinator	Subordinators are divided into different types and all other subordinators are grouped into an 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.
Title	Deleted	No applicable broader category.
Verb-ING (not in standard progressive form)	Deleted	It could be an auxiliary omission and thus in progressive form or it could be a gerund. Rather than check each instance to clarify and rather than group these instances into either general verbs or general nouns and misclassify some, it was decided to just delete the feature altogether from the final feature set.
Verb+that complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.
WH word + contracted verb	Pooled with pronoun + contracted verb in a ‘Contracted forms’ variable.	I could have either deleted this feature or combined with pronoun with contracted verb to a broader category of contracted forms. I tested both conditions by running two different MCAs: one where it was deleted and the other where it was combined with pronoun with contracted verb. I correlated the coordinates of the individual tweets for the first 10 dimensions from both sets of results and this revealed that they were strongly correlated, suggesting that there was little effect by pooling. I also compared the contribution and coordinate of ‘contracted forms’ and ‘pronoun with contracted verb’ in each analysis on each dimension to observe if the pooling led to any substantial difference. There was no substantial difference. One of the benefits from including this feature by pooling it was an increase in percentage of explained variance from the eigenvalues. It was therefore decided that the feature would be pooled.
WH-clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.

WH-word + BE	Deleted	Not enough to combine with other question features.
WH-word + DO	Deleted	Not enough to combine with other question features.

After this pooling process was completed, each general English tweet was analysed for the presence or absence of the following 63 linguistic features presented in Table 50.

Table 50: The feature set used in MDA of General Englihs Tweets

General Twitter Feature Set	**Feature Description (incl. Pooled Features)**
Amplifier	Refers to adverbs used to intensify the verb/adjective
Analytic_Negation	Refers to 'not' plus contracted forms
Attributive_Adjective	Adjectives that come before the noun and any other adjective not tagged as predicative.
Auxiliary_DO	Refers to any form of DO that is followed by (up to three adverbs and) a verb.
Capitalisation	Refers to two or more capital letters that is not tagged as an acronym/ URL/ mentioned username
Colon	Refers to the use of colons
Comma	Refers to the use of commas
Complementation	Verb+*that* complement clause, Noun+*that* complement clause, Adjective+*that* complement clause, Adjective+*to* complement clause, WH-clause.
Contracted_Forms	Refers to when a pronoun has the verb contracted and when the WH word has the verb contracted
Contrastive_Conjunct	Refers to conjunctions that signal a contrast is being made
Coordinating_Conjunct	Refers to coordinating conjunctions.
Definite_Article	Refers to the use of the definite article
Demonstrative_Determiner	Refers to *this, that, these, those* followed by a noun (which can be preceded by adjectives, adverbs).
Demonstrative_Pronoun	Refers to the use of this, that, these, those as a pronoun; that is NOT followed by noun
Ellipsis	Refers to three or more full stops

Emoticon/Emoji	Refers to anything tagged by the Gimpel et al. (2011) tagger as an emoticon and some unicodes.
Exclamation_Mark	Refers to the use of exclamation marks
First_Person_Pronoun	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the first person: singular and plural plus contracted forms
Full_Stop	Refers to use of full stop
General_Subordinator	Refers to all subordinators, including those indicating time, place, cause, condition, concession.
Gradation	Refers to adjectives and nouns in superlative and comparative forms.
Hashtag	Refers to the use of Hashtag
HAVE_Main_Verb	Refers to when any form of HAVE is the main verb
Imperative	Refers to clauses in imperative mood
Indefinite_Article	Refers to use of indefinite article
Indefinite_Pronoun	Refers to pronouns which indicate quantity or are indefinite pronouns
Infinitive	Refers to verbs in infinitive form that is not adjective + to complement clause. Also refers to split infinitives: verbs in infinitives form separated by adverb(s).
Initial_Mention	Refers to Tweet initial mentioning
Interjection	Laughter, Negative Interjection, Positive Interjection, anything tagged as interjection by Twitter tagger.
IT	Refers to any form of pronoun IT: contracted, reflexive, possessive and possessive determiner
Modal_Possibility	Refers to modals indicating probability/possibility/ability
Modal_Prediction	Refers to modals indicating prediction and BE+going to construction
Nominalisation	Refers to when verbs/adjectives are converted into nouns
Non-Initial_Mention	Refers to mentioning that is not initial
Numeral_Determiner	Refers to use of numerals functioning as determiners
Numeral_Noun	Refers to use of numerals functioning as nouns
Object_Pronoun	Refers to use of pronouns in their objective form
Other_Adverb	Refers to other adverbs that are not tagged as amplifiers, downtoners, time and place adverbials,

	quantifying adverbs, adverbs of usuality. However, Downtoner, Place Adverb, Quantifying Adverb, Adverbs of Frequency/Usuality are pooled.
Other_Noun	Refers to other nouns that are not tagged as numeral, quantifiers, nominalisations, ordinals.
Other_Verb	Suasive verbs: Refers to verbs which refer to persuasion and Pro-verb DO: Refers to DO used as a main verb
Passive	Agentless- and By- Passives
Past_Tense_Verb	Refers to verbs in their past tense form that are not in perfect aspect
Perception_Verb	Refers to verbs of perception
Phrasal_Verb	Refers to both prepositional and particle verbs
Possession	Refers to determiners, pronouns, proper nouns, and nouns which indicate possession
Predicative_Adjective	Refers to adjectives which come after a copular verb
Preposition	Refers to the use of prepositions
Private_Verb	Refers to private verbs: used to encode feelings, opinions, emotions, cognition
Profanity	Refers to words that can be used to offend/abuse as well as swear words generally. They may also be used harmlessly
Progressive	Refers to any form of BE plus (up to 2/3 adverbs and) verb ending in -ING
Proper_Noun	Refers to anything tagged as a proper noun
Public_Verb	Refers to public verbs: used to report on speech
Quantifier_Determiner	Refers to quantifiers used as a determiner
Question_Mark	Refers to the use of question mark
Second_Person_Pronoun	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the second person: singular and plural plus contracted forms
Stance_Verb	Refers to verbs used to encode stance
Stative_Form	Refers to when BE is the main verb and when BE is in its copular form; that is, when it is followed by a predicative adjective. Also it refers to the use of there in its existential form and thus not as a place adverb, and also includes other copular verbs in their copula form: followed by predicative adjective.
Subject_Pronoun	Refers to pronouns in their subject form

Third_Person_Pronoun	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the third person: singular and plural plus contracted forms
Third_Person_Singular_Verb	Refers to verbs ending in -s
Time_Adverb	Refers to adverbs indicating time
URL	Refers to URLs: can be meme, gif, status, link to website, video etc.
WH-Word	Refers to use of WH words

Similar to general English Twitter described above, the frequencies of occurrence of each linguistic feature in the trolling corpus were computed and the features that occurred in fewer than five percent of the trolling tweets (*n* = 209) were either pooled with other similar features into a broader linguistic category, combined with a broader 'other' variable or they were deleted from the final feature set. Table 51 presents the features occurring in fewer than five percent of trolling tweets and the decision that was made with respect to pooling or deleting features and the justification for these decisions.

Table 51: The features occurring in fewer than five percent of the Trolling tweets and the decision and justifcation for inclusion/exclusion in the final feature set

Features < 5 % of Trolling Tweets	**Decision**	**Justification**
Acronym	Deleted	Acronyms are already counted as the part of speech tag assigned (e.g. U.S.A_^ is assigned proper noun tag)
Adj+*that* complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.
Adj+*to* complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.

Adverbs of frequency/usuality	Pooled with General Adverbs	Adverbs are divided into different types and all other adverbs not specified are grouped into a broader 'other adverbs' category. If the specific type of adverb occurs infrequently then each instance can be re-combined with the 'other adverb' feature.
Bracket	Deleted	No applicable broader category, albeit 'punctuation'; however this would lead to the loss of the distinction between question marks and exclamation marks.
By-Passive	Pooled with Agentless Passives into passive constructions	Passive constructions were divided into different types and therefore re-combined to form the broader category.
Cause Subordinator	Pooled with general subordinator	Subordinators are divided into different types and all other subordinators are grouped into an 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.
Colon	Deleted	No applicable broader category, albeit 'punctuation'; however this would lead to the loss of the distinction between question marks and exclamation marks.
Concessive Subordinator	Pooled with general subordinator	Subordinators are divided into different types and all other subordinators are grouped into an 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.
Copular verbs (not BE)	Pooled with BE as a main verb	I could have deleted this feature from the feature set, or combined either with ‘General/Other Verbs’ or pooled with BE as a main verb and existential there into a category of Stative forms. I tested all by running the analysis on all types and then correlated the coordinates of the individual tweets from the results. I also compared the contribution and coordinate of ‘Stative forms’ and ‘BE as a main verb’, as well as ‘General/Other verbs’ pre- and post- pooling in each analysis on each dimension to observe if the pooling led to any substantial difference.
Downtoner Adverb	Pooled with General Adverbs	Adverbs are divided into different types and all other adverbs not specified are grouped into a broader 'other adverbs' category. If the specific type of adverb occurs infrequently then each instance can be re-combined with the 'other adverb' feature.

Emoji/Emoticon	Deleted	No applicable broader category, albeit 'CMC features' but this would lose the different positions of mentioning distinctions which do occur in more than five percent of the tweets.
Existential *there*	Pooled with BE as a main verb and copular verbs	I could have deleted or combined with BE as a main verb and copular verbs into a category of 'stative forms'. I tested both by running the analysis on both scenarios and then correlated the coordinates of the individual tweets from the results on all the dimensions. I also compared the contribution and coordinate of 'Stative forms' and 'BE as a main verb', in each analysis on each dimension to observe if the pooling led to any substantial difference.
Initial verb	Deleted	Whatever the verb is, it would also be classified as either one of the verb types or in the 'other verb' category. Therefore it does not need to be pooled with broader verb category. I could have combined with other initial verbs. However, I tested this by running the analysis on the feature combined with other initial verbs as well as with this feature deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb -S	Deleted	This feature is already counted as third person singular verb form regardless of initial position. I could have combined it with other initial verbs. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb BE	Deleted	I could have combined into a broader category of initial verbs with other initial verb instances. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb DO	Deleted	I could have combined into a broader category of initial verbs with other initial verb instances. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.

Initial verb HAVE	Deleted	I could have combined into a broader category of initial verbs with other initial verb instances. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb ING	Deleted	It could be an auxiliary omission and thus in progressive form or it could be a gerund. Rather than check each instance to clarify and rather than group these instances into either general verbs or general nouns and misclassify some, it was decided to just delete the feature altogether from the feature set. I could have combined with initial verbs (but it might not have been one). However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb MODAL	Deleted	The type of modal is counted any way regardless of whether it is positioned initially. I could have combined it with other initial verbs. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb PAST	Deleted	This feature is already counted as past tense verb regardless of initial position. I could have combined it with other initial verbs. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Initial verb QUESTION	Deleted	Too few instances to combine with other question features. I could have combined with other initial verb types. However, I tested this by running the MCA on one feature set where the feature combined all initial verbs, as well as another feature set where this feature was deleted. Overall, the new initial verb feature influenced the dimensions too substantially and made the dimensions far less interpretable.
Laughter Interjection	Pooled with General Interjections	All specific types of interjections occurred in fewer than five percent of tweets. These specific types were therefore combined into a broader category of interjections.

Modal of Necessity	Deleted	I could have combined all modals together but this would lose the distinction between possibility and prediction modals, which contribute differently on different dimensions.
Negative Interjection	Pooled with General Interjections	All specific types of interjections occurred in fewer than five percent of tweets. These specific types were therefore combined into a broader category of interjections.
Noun+*that* complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.
Ordinal determiner	Deleted	Not enough instances of either kind of ordinal to combine to create an ordinal feature. I could have combined with other determiners (e.g. definite and indefinite article and demonstratives and possessive) but this would lose the distinction between them.
Ordinal Noun	Deleted	Not enough instances of either kind of ordinal to combine to create an ordinal feature. I could have combined to general noun category, but because I deleted ordinal determiners from the feature set I also decided to delete ordinal nouns.
Other Conjunctions	Deleted	Too few instances to combine with other frequent and more specific kinds of conjunctions.
Pied-piping relative	Pooled with Relatives	Relative clauses were divided into different types and therefore re-combined to form broader category of relatives.
Place subordinator	Pooled with general subordinator	Subordinators are divided into different types and all other subordinators are grouped into a 'other subordinator' feature. Therefore, if a specific type does not occur frequently it can re-join the 'other subordinator' feature category.
Positive Interjection	Pooled with General Interjections	All specific types of interjections occurred in fewer than five percent of tweets. These specific types were therefore combined into a broader category of interjections.
Possessive Noun	Pooled with Possession	The only feature denoting possession that occurred in more than 5% of the tweets was possessive determiner. These features were combined to create one variable of possession.
Possessive Pronoun	Pooled with Possession	The only feature denoting possession that occurred in more than 5% of the tweets was possessive determiner. These features were combined to create one variable of possession. The type of possessive pronoun is also counted (e.g. first, second, third).

Possessive Proper Noun	Pooled with Possession	The only feature denoting possession that occurred in more than 5% of the tweets was possessive determiner. These features were combined to create one variable of possession.
Pre-Determiner	Deleted	Too few instances to combine with quantifying pre-determiners to create general pre-determiner category.
Quantifying Adverb	Pooled with General Adverbs	Adverbs were divided into different types and all other adverbs that do not fall in these particular categories are grouped into a category called ‘other adverbs’. Therefore, if a specific type does not occur frequently it can rejoin the ‘other adverbs’ category. Quantifying adverbs could have also been grouped with other quantifiers of different parts-of-speech (e.g. quantifying -determiners, -pre-determiners, -pronouns). They were not grouped this way because quantifying determiners and pronouns occurred in more than five percent of trolling tweets and I did not want to lose this part-of-speech distinction by grouping them all together.
Quantifying Pre-Determiner	Deleted	Too few instances to combine with other pre-determiners to create general pre-determiner category.
Reflexive Pronoun	Deleted	No applicable broader category, albeit 'pronouns'. Reflexive pronouns are counted depending on first, second or third person, or IT.
Relative Clause Object Gap	Pooled with Relatives	Relative clauses were divided into different types and therefore re-combined to form broader category of relatives.
Semi Colon	Deleted	No applicable broader category, albeit 'punctuation'; however this would lead to the loss of the distinction between question marks and exclamation marks.
Split Infinitive	Pooled with infinitives	Split infinitives were separated from infinitives as a particular type and so therefore were recombined to the broader category.
Suasive verb	Pooled with General verbs	Different types of verbs were distinguished from general verbs and therefore infrequent types can be recombined with broader verb category.
Subordinator with ellipted subject	Deleted	No applicable broader category. If it is a specific type of subordinator it will be classified as such as well.

Superlative	Pooled with comparatives to form ‘Gradation’	I could have deleted or like the general twitter study I could have combined with comparatives to form a broader 'gradation' category. I tested this by running separate MCAs on a feature set with only comparatives and another one with combined superlatives and comparatives. Both results were strongly positively correlated. I combined the features in order to make sure superlatives were not discounted.
Title	Deleted	No applicable broader category.
Verb-ING (not progressive)	Deleted	It could be an auxiliary omission and thus in progressive form or it could be a gerund. Rather than check each instance to clarify and rather than group these instances into either general verbs or general nouns and misclassify some, it was decided to just delete the feature altogether from the final feature set.
Verb+*that* complement clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.
WH word + contracted verb	Pooled with Pronoun + contracted verb in a ‘Contracted forms’ variable	I could have either deleted this feature or combined with pronoun with contracted verb to a broader category of contracted forms. I tested both conditions by running two different MCAs: one where it was deleted and the other where it was combined with pronoun with contracted verb. I correlated the coordinates of the individual tweets for the first 10 dimensions from both sets of results and this revealed that they were strongly correlated, suggesting that there was little effect by pooling. I also compared the contribution and coordinate of ‘contracted forms’ and ‘pronoun with contracted verb’ in each analysis on each dimension to observe if the pooling led to any substantial difference. There was no substantial difference. One of the benefits from including this feature by pooling it was an increase in percentage of explained variance from the eigenvalues. It was therefore decided that the feature would be pooled.
WH-clause	Pooled with Complementation	All specific types of complement clauses occurred in fewer than five percent of tweets. These specific types were therefore combined to form one broad category of complementation.
WH-word + BE	Deleted	Not enough to combine with other question features.
WH-word + DO	Deleted	Not enough to combine with other question features.

After this pooling and deleting process, each trolling tweet was analysed for the presence or absence of the following 69 linguistic features presented in Table 52.

Table 52: The feature set used in the MDA of Twitter Trolling

Trolling Feature Set	**Feature Description and Pooled Features**
Amplifier	Refers to adverbs used to intensify the verb/adjective
Analytic_Negation	Refers to 'not' plus contracted forms
Attributive_Adjective	Adjectives that come before the noun and any other adjective not tagged as predicative.
Auxiliary_DO	Refers to any form of DO that is followed by (up to three adverbs and) a verb.
Capitalisation	Refers to two or more capital letters that is not tagged as an acronym/ URL/ mentioned username
Comma	Refers to the use of commas
Complementation	Verb+*that* complement clause, Noun+*that* complement clause, Adjective+*that* complement clause,Adjective+*to* complement clause, WH-clause.
Conditional_Subordinator	Refers to subordinators indicating a condition
Contracted_Forms	Refers to when a pronoun has the verb contracted and when the WH word has the verb contracted
Contrastive_Conjunct	Refers to conjunctions that signal a contrast is being made
Coordinating_Conjunct	Refers to coordinating conjunctions.
Definite_Article	Refers to the use of the definite article
Demonstrative_Determiner	Refers to *this, that, these, those* followed by a noun (which can be preceded by adjectives, adverbs).
Demonstrative_Pronoun	Refers to the use of this, that, these, those as a pronoun; that is NOT followed by noun
Ellipsis	Refers to three or more full stops
Exclamation_Mark	Refers to the use of exclamation marks
First_Person_Pronoun	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the first person: singular and plural plus contracted forms
Full_Stop	Refers to use of full stop

Gerund	Refers to prepositional complement: when a preposition is followed by noun in -ing form (but this is tagged by Gimpel tagger as a verb)
Gradation	Refers to adjectives and nouns in superlative and comparative forms.
Hashtag	Refers to the use of Hashtag
HAVE_Main_Verb	Refers to when any form of HAVE is the main verb
Imperative	Refers to clauses in imperative mood
Indefinite_Article	Refers to use of indefinite article
Indefinite_Pronoun	Refers to pronouns which indicate quantity or are indefinite pronouns
Infinitive	Refers to verbs in infinitive form that is not adjective + to complement clause. Also refers to split infinitives: verbs in infinitives form separated by adverb(s).
Initial_Mention	Refers to Tweet initial mentioning
Interjection	Laughter, Negative Interjection, Positive Interjection, anything tagged as interjection by Twitter tagger.
IT	Refers to any form of pronoun IT: contracted, reflexive, possessive and possessive determiner
Modal_Possibility	Refers to modals indicating probability/possibility/ability
Modal_Prediction	Refers to modals indicating prediction and BE+going to construction
Nominalisation	Refers to when verbs/adjectives are converted into nouns
Non-Initial_Mention	Refers to mentioning that is not initial
Numeral_Determiner	Refers to use of numerals functioning as determiners
Numeral_Noun	Refers to use of numerals functioning as nouns
Object_Pronoun	Refers to use of pronouns in their objective form
Other_Adverb	Refers to other adverbs that are not tagged as amplifiers, downtoners, time and place adverbials, quantifying adverbs, adverbs of usuality. However, downtoner, Quantifying Adverb, and adverbs of frequency/Usuality are pooled.
Other_Noun	Refers to other nouns that are not tagged as numeral, quantifiers, nominalisations, ordinals.
Other_Subordinator	Refers to any subordinator that is not time subordinator, including Place subordinator, Cause subordinator, and Concessive subordinator.

Other_Verb	Suasive verbs: Refers to verbs which refer to persuasion
Passive	Agentless- and By- Passives: Refers to use of passive voice with and without the inclusion of an agent in a by clause
Past_Tense_Verb	Refers to verbs in their past tense form that are not in perfect aspect
Perception_Verb	Refers to verbs of perception
Perfect_Aspect	Refers to any form of HAVE + verb in past participle form
Phrasal_Verb	Refers to both prepositional and particle verbs
Place_Adverb	Refers to adverbs indicating place
Possession	Refers to determiners, pronouns, proper nouns, and nouns which indicate possession
Predicative_Adjective	Refers to adjectives which come after a copular verb
Preposition	Refers to the use of prepositions
Private_Verb	Refers to private verbs: used to encode feelings, opinions, emotions, cognition
Pro-Verb_DO	Refers to DO used as a main verb
Profanity	Refers to words that can be used to offend/abuse as well as swear words generally. They may also be used harmlessly
Progressive	Refers to any form of BE plus (up to 2/3 adverbs and) verb ending in -ING
Proper_Noun	Refers to anything tagged as a proper noun
Public_Verb	Refers to public verbs: used to report on speech
Quantifier_Determiner	Refers to quantifiers used as a determiner
Question_Mark	Refers to the use of question mark
Relatives	Refers to the use of relative clauses with subject/object gap, and pied-piping relative, which refers to the use of preposition + relative pronoun used in order to avoid stranded preposition.
Second_Person_Pronoun	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the second person: singular and plural plus contracted forms
Stance_Verb	Refers to verbs used to encode stance

Stative_Form	Refers to when BE is the main verb and when BE is in its copular form; that is, when it is followed by a predicative adjective. Also it refers to the use of there in its existential form and thus not as a place adverb, and also includes other copular verbs in their copula form: followed by predicative adjective.
Subject_Pronoun	Refers to pronouns in their subject form
Synthetic_Negation	Refers to use of nor, neither and no - but not as interjection
Third_Person_Pronoun	Refers to pronouns: subject/object/possessive/reflexive and possessive determiners that refer to the third person: singular and plural plus contracted forms
Third_Person_Singular_Verb	Refers to verbs ending in -s
Time_Adverb	Refers to adverbs indicating time
Time_Subordinator	Refers to subordinators indicating time
URL	Refers to URLs: can be meme, gif, status, link to website, video etc.
WH-Word	Refers to use of WH words

Appendix 2: Correlation Matrices of Samples of Corpora.

Table 53: Correlation Matrices of the Coordinates and Contributions from MDAs of Samples of General English Twitter

	Correlation_of_Coordinates_500-word_sample_1_and_2				
	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9568	0.14	0.0093	-0.114	-0.013
Dim_2	-0.0182	-0.855	0.0146	0.129	-0.047
Dim_3	0.0472	0.036	0.6637	-0.031	-0.068
Dim_4	0.0222	0.036	0.0215	0.162	-0.184
Dim_5	-0.0033	-0.106	-0.2653	0.109	0.303
	Correlation_of_Contributions_500-word_sample_1_and_2				
	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.888	0.0439	0.335	0.323	0.294
Dim_2	-0.036	0.9215	0.152	0.1842	-0.011
Dim_3	0.088	0.1527	0.548	-0.0421	0.372
Dim_4	0.15	-0.0074	0.035	0.1253	0.027
Dim_5	0.012	-0.0986	-0.033	0.0097	0.283
	Correlation_of_Coordinates_500-word_sample_3_and_4				
	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.951	0.1685	-0.095	-0.051	0.092
Dim_2	-0.057	-0.9225	0.108	0.137	0.019
Dim_3	-0.097	0.0051	0.276	-0.402	0.056
Dim_4	0.132	0.0099	0.219	0.547	0.194
Dim_5	-0.026	-0.0393	-0.324	0.204	0.232
	Correlation_of_Contributions_500-word_sample_3_and_4				
	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.913	-0.048	0.34	0.155	0.322
Dim_2	-0.029	0.922	0.07	0.125	0.065
Dim_3	0.328	0.03	0.16	0.434	0.205
Dim_4	0.018	0.292	0.29	0.308	0.145
Dim_5	0.073	0.089	0.42	-0.024	0.142
	Correlation_of_Coordinates_500-word_sample_5_and_6				
	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.969	-0.0086	0.014	-0.0646	-0.058
Dim_2	0.248	-0.9119	0.26	0.0156	-0.067
Dim_3	0.134	-0.1537	-0.175	-0.5814	0.159
Dim_4	-0.061	-0.0271	-0.399	0.0089	0.212
Dim_5	-0.125	0.1538	0.114	-0.0362	0.077

Correlation_of_Contributions_500-word_sample_5_and_6

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.936	-0.0384	0.36	0.17	-0.151
Dim_2	0.018	0.8624	0.11	0.13	0.138
Dim_3	0.357	0.0315	0.39	0.69	0.101
Dim_4	0.232	0.0281	0.39	0.22	0.074
Dim_5	0.091	-0.0069	0.06	0.54	0.296

Correlation_of_Coordinates_500-word_sample_7_and_8

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.961	0.1555	0.011	0.2497	0.11
Dim_2	-0.057	-0.8541	-0.072	-0.1491	-0.001
Dim_3	0.021	-0.0035	0.373	-0.0426	0.146
Dim_4	0.103	0.1705	-0.457	0.1727	0.037
Dim_5	-0.042	0.0436	0.168	0.0053	-0.092

Correlation_of_Contributions_500-word_sample_7_and_8

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.902	-0.048	0.29	0.18	-0.011
Dim_2	-0.044	0.916	0.16	-0.1	0.116
Dim_3	0.195	0.286	0.13	0.31	0.112
Dim_4	0.176	0.023	0.62	0.1	0.16
Dim_5	0.226	0.173	0.12	0.24	0.141

Correlation_of_Coordinates_500-word_sample_9_and_10

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.966	-0.15	-0.084	0.01	0.041
Dim_2	-0.127	0.881	-0.123	0.1	-0.036
Dim_3	0.03	0.059	0.445	-0.32	-0.075
Dim_4	-0.05	0.055	0.032	0.32	-0.26
Dim_5	0.172	-0.182	-0.134	0.23	0.162

Correlation_of_Contributions_500-word_sample_9_and_10

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.899	-0.051	0.391	0.444	0.327
Dim_2	-0.046	0.897	0.152	-0.054	0.085
Dim_3	0.238	0.289	0.698	0.451	0.476
Dim_4	0.318	-0.017	0.318	0.575	0.146
Dim_5	0.067	0.039	0.066	0.148	0.348

Correlation_of_Coordinates_1000-word_sample_11_and_12

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.973	0.215	0.02236	-0.062	0.018
Dim_2	-0.099	-0.949	-0.00014	0.097	0.015
Dim_3	0.187	0.087	0.66795	-0.421	0.08
Dim_4	-0.151	-0.103	0.32998	0.57	0.15
Dim_5	-0.085	-0.01	-0.17712	-0.005	0.443

Correlation_of_Contributions_1000-word_sample_11_and_12

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.95	-0.066	0.4185	0.13	0.0036
Dim_2	-0.04	0.947	-0.0033	0.14	-0.0056

Dim_3	0.42	-0.017	0.7555	0.35	0.1144
Dim_4	0.14	0.135	0.2368	0.6	0.2634
Dim_5	-0.02	0.478	-0.0374	0.15	0.3472

Correlation_of_Coordinates_1000-word_sample_13_and_14

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.975	-0.0873	-0.024	0.1	0.038
Dim_2	-0.14	0.9437	0.017	-0.17	-0.103
Dim_3	-0.012	-0.0783	-0.338	-0.4	0.386
Dim_4	0.107	-0.0011	-0.386	0.68	0.307
Dim_5	0.128	-0.0165	0.445	0.16	0.24

Correlation_of_Contributions_1000-word_sample_13_and_14

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.933	-0.0651	0.17	0.439	0.089
Dim_2	-0.083	0.9634	0.41	-0.035	0.057
Dim_3	0.363	-0.004	0.64	0.399	0.449
Dim_4	0.258	-0.0329	0.26	0.6	0.394
Dim_5	0.215	0.2652	0.23	0.204	0.054

Correlation_of_Coordinates_1000-word_sample_15_and_16

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	-0.156	0.0147	0.116	0.073
Dim_2	-0.086	0.909	0.00051	-0.049	-0.059
Dim_3	-0.076	0.141	-0.10145	-0.57	0.445
Dim_4	-0.021	0.079	0.66715	-0.061	-0.138
Dim_5	0.166	-0.106	0.27121	0.433	0.398

Correlation_of_Contributions_1000-word_sample_15_and_16

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9678	-0.0076	0.281	0.163	0.1
Dim_2	-0.0019	0.9122	0.142	0.054	0.16
Dim_3	0.2756	0.3271	0.338	0.375	0.49
Dim_4	0.1804	0.0104	0.584	0.155	0.18
Dim_5	0.0926	0.1051	-0.051	0.566	0.4

Correlation_of_Coordinates_1000-word_sample_17_and_18

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9643	-0.2001	0.0028	0.124	0.022
Dim_2	-0.0082	0.9466	-0.064	-0.046	0.108
Dim_3	0.0349	0.035	0.4933	-0.039	0.031
Dim_4	0.0032	0.0077	-0.1552	-0.548	0.126
Dim_5	0.0345	-0.1966	-0.3547	-0.06	0.114

Correlation_of_Contributions_1000-word_sample_17_and_18

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9469	-0.095	0.3	0.4	-0.0065
Dim_2	-0.0311	0.945	0.17	0.1	0.2657
Dim_3	-0.0027	0.583	0.54	-0.02	0.2601
Dim_4	0.3128	0.099	0.11	0.64	0.1611
Dim_5	0.2504	0.051	0.42	0.15	0.2383

Correlation_of_Coordinates_1000-word_sample_19_and_20

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.986	-0.049	0.088	0.15	-0.0047
Dim_2	-0.161	0.967	-0.056	-0.1	-0.0126
Dim_3	0.083	-0.107	0.341	-0.57	-0.0488
Dim_4	-0.034	0.013	-0.123	-0.2	0.4917
Dim_5	0.051	0.031	0.276	0.47	0.5156

Correlation_of_Contributions_1000-word_sample_19_and_20

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.972	-0.117	0.248	0.221	0.028
Dim_2	-0.055	0.972	-0.034	0.021	0.111
Dim_3	0.277	0.074	0.572	0.312	-0.058
Dim_4	0.295	0.125	0.14	0.294	0.357
Dim_5	-0.031	0.076	0.142	0.349	0.463

Correlation_of_Coordinates_2000-word_sample_21_and_22

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9898	-0.081	-0.067	0.11	0.091
Dim_2	-0.1492	0.969	0.186	-0.1	-0.036
Dim_3	-0.0023	-0.043	0.336	0.66	-0.11
Dim_4	-0.0694	-0.052	0.785	-0.48	-0.136
Dim_5	0.0826	0.002	-0.106	0.17	0.49

Correlation_of_Contributions_2000-word_sample_21_and_22

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.969	-0.045	0.15	0.3758	0.147
Dim_2	-0.082	0.952	0.13	-0.0408	0.334
Dim_3	0.456	-0.023	0.24	0.8272	-0.024
Dim_4	0.31	0.158	0.66	0.4507	0.157
Dim_5	0.028	0.241	0.13	0.0031	0.207

Correlation_of_Coordinates_2000-word_sample_23_and_24

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9884	-0.074	0.091	-0.0326	-0.04
Dim_2	-0.1641	0.976	-0.01	0.002	0.095
Dim_3	0.0032	0.035	0.502	0.597	-0.166
Dim_4	-0.1103	0.076	-0.421	0.5812	0.453
Dim_5	0.0687	-0.126	0.554	0.0104	0.288

Correlation_of_Contributions_2000-word_sample_23_and_24

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.971	-0.06	0.317	0.12	-0.062
Dim_2	-0.087	0.975	-0.034	0.3	0.101
Dim_3	0.391	0.018	0.452	0.69	0.014
Dim_4	0.302	0.168	0.666	0.21	0.289
Dim_5	-0.039	-0.01	0.221	0.24	0.193

Correlation_of_Coordinates_2000-word_sample_25_and_26

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.986	-0.08409	-0.0063	-0.09	-0.027
Dim_2	-0.144	0.9646	0.1298	0.055	0.078
Dim_3	0.087	-0.00046	0.5572	-0.628	0.177

Dim_4	-0.124	0.00566	0.4239	0.766	0.067
Dim_5	0.148	-0.07319	-0.1934	-0.082	0.782

Correlation_of_Contributions_2000-word_sample_25_and_26

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.974	-0.054	0.228	0.392	-0.094
Dim_2	-0.014	0.96	0.094	0.063	0.102
Dim_3	0.451	0.143	0.55	0.646	0.047
Dim_4	0.273	0.303	0.544	0.635	0.193
Dim_5	0.201	0.08	0.258	0.214	0.73

Correlation_of_Coordinates_2000-word_sample_27_and_28

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.992	0.136	-0.065	-0.104	0.017
Dim_2	-0.099	-0.967	-0.024	0.116	0.077
Dim_3	0.054	-0.035	0.823	-0.102	-0.034
Dim_4	-0.106	0.03	-0.193	0.758	-0.182
Dim_5	0.068	0.105	-0.04	-0.057	0.51

Correlation_of_Contributions_2000-word_sample_27_and_28

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.981	-0.045	0.321	0.3	0.051
Dim_2	-0.061	0.973	-0.044	0.07	0.227
Dim_3	0.321	0.124	0.881	0.25	0.288
Dim_4	0.24	0.314	0.393	0.71	0.143
Dim_5	0.106	0.07	0.102	0.33	0.488

Correlation_of_Coordinates_2000-word_sample_29_and_30

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.989	-0.145	0.064	0.01	-0.0075
Dim_2	0.125	-0.969	0.039	-0.076	0.0119
Dim_3	0.061	0.052	0.705	0.376	0.1074
Dim_4	-0.058	0.058	-0.637	0.594	0.0685
Dim_5	0.144	-0.233	0.115	-0.163	0.2208

Correlation_of_Contributions_2000-word_sample_29_and_30

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.975	-0.05	0.36	0.17	-0.03
Dim_2	-0.043	0.978	-0.03	0.27	0.075
Dim_3	0.426	-0.012	0.78	0.28	-0.013
Dim_4	0.151	0.414	0.45	0.62	0.13
Dim_5	0.179	0.271	0.07	0.3	0.127

Correlation_of_Coordinates_3000-word_sample_31_and_32

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.993	-0.1038	0.141	-0.035	0.062
Dim_2	-0.1504	0.9808	-0.036	0.163	-0.087
Dim_3	-0.0062	0.0047	-0.041	0.871	0.025
Dim_4	-0.0337	-0.0028	-0.885	-0.032	-0.125
Dim_5	0.0567	0.0285	-0.026	0.087	0.723

Correlation_of_Contributions_3000-word_sample_31_and_32

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5

Dim_1	0.984	-0.053	0.3626	0.164	0.086
Dim_2	-0.093	0.977	-0.0323	0.128	0.182
Dim_3	0.218	0.189	-0.0084	0.904	0.106
Dim_4	0.367	0.023	0.9362	0.065	0.096
Dim_5	-0.055	0.086	0.0452	0.229	0.681

Correlation_of_Coordinates_3000-word_sample_33_and_34

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.991	-0.128	-0.075	0.014	0.0384
Dim_2	-0.112	0.98	0.073	-0.092	0.002
Dim_3	0.057	0.025	-0.677	0.674	-0.0344
Dim_4	-0.054	0.131	0.716	0.514	0.0726
Dim_5	0.142	-0.027	-0.205	0.071	0.5685

Correlation_of_Contributions_3000-word_sample_33_and_34

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.983	-0.093	0.309	0.243	-0.122
Dim_2	-0.018	0.979	0.047	0.165	0.2
Dim_3	0.364	-0.045	0.77	0.401	-0.063
Dim_4	0.116	0.317	0.362	0.551	0.288
Dim_5	0.144	0.144	0.256	0.085	0.499

Correlation_of_Coordinates_3000-word_sample_35_and_36

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.993	-0.097	-0.101	-0.0203	-0.027
Dim_2	-0.13	0.984	0.084	0.1415	0.037
Dim_3	0.035	0.005	-0.748	0.4799	-0.071
Dim_4	-0.056	0.015	0.53	0.6798	0.138
Dim_5	0.11	-0.081	-0.042	0.0071	0.463

Correlation_of_Contributions_3000-word_sample_35_and_36

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9842	-0.042	0.311	0.15	0.028
Dim_2	-0.0671	0.983	0.044	0.25	0.056
Dim_3	0.3868	-0.047	0.808	0.27	0.02
Dim_4	0.199	0.338	0.312	0.71	0.156
Dim_5	0.0051	0.053	0.023	0.12	0.289

Correlation_of_Coordinates_3000-word_sample_37_and_38

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.993	-0.141	-0.027	-0.093	0.131
Dim_2	-0.108	0.983	0.091	0.034	-0.023
Dim_3	-0.079	0.092	0.215	0.85	-0.263
Dim_4	-0.074	0.058	0.881	-0.151	0.018
Dim_5	-0.016	-0.011	0.244	0.085	0.462

Correlation_of_Contributions_3000-word_sample_37_and_38

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.983	-0.0584	0.181	0.3647	0.018
Dim_2	-0.064	0.9812	0.199	-0.0063	0.245
Dim_3	0.327	0.0615	0.061	0.9233	0.133
Dim_4	0.172	0.1506	0.901	0.0969	0.063

Dim_5	-0.048	0.0059	0.075	0.0393	0.298

Correlation_of_Coordinates_3000-word_sample_39_and_30

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.993	-0.106	0.105	-0.037	-0.02
Dim_2	-0.153	0.971	-0.021	0.087	0.077
Dim_3	0.083	-0.105	0.921	-0.111	-0.044
Dim_4	0.091	-0.04	-0.011	-0.792	0.242
Dim_5	-0.089	0.044	-0.075	0.311	0.569

Correlation_of_Contributions_3000-word_sample_39_and_30

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	-0.079	0.352	0.159	-0.085
Dim_2	-0.023	0.972	0.018	0.325	0.233
Dim_3	0.387	-0.057	0.96	0.033	-0.025
Dim_4	0.13	0.152	-0.04	0.874	0.136
Dim_5	-0.096	0.074	0.047	0.135	0.78

Correlation_of_Coordinates_4000-word_sample_41_and_42

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.995	0.154	0.0045	-0.092	-0.013
Dim_2	-0.098	-0.985	0.0895	0.088	0.04
Dim_3	-0.056	-0.059	-0.1807	0.912	0.188
Dim_4	-0.037	-0.026	0.9092	-0.017	0.357
Dim_5	0.099	-0.013	-0.1415	-0.179	0.707

Correlation_of_Contributions_4000-word_sample_41_and_42

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.989	-0.053	0.28	0.362	-0.054
Dim_2	-0.059	0.987	0.1	0.09	0.021
Dim_3	0.321	0.095	0.22	0.942	0.016
Dim_4	0.149	0.175	0.87	0.168	0.225
Dim_5	0.012	0.089	0.23	0.151	0.761

Correlation_of_Coordinates_4000-word_sample_43_and_44

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.996	-0.114	-0.05	-0.101	0.022
Dim_2	-0.109	0.985	0.1211	0.027	0.061
Dim_3	0.078	-0.076	0.1655	-0.895	0.05
Dim_4	-0.089	0.042	0.8723	0.321	0.113
Dim_5	0.047	-0.012	0.0055	0.11	0.775

Correlation_of_Contributions_4000-word_sample_43_and_44

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	-0.075	0.154	0.35787	-0.081
Dim_2	-0.042	0.984	0.174	-0.00017	0.028
Dim_3	0.395	-0.059	0.033	0.91722	0.026
Dim_4	0.161	0.338	0.928	0.05335	0.063
Dim_5	-0.027	0.105	0.168	0.00075	0.794

Correlation_of_Coordinates_4000-word_sample_45_and_46

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9935	0.129	-0.12	0.0022	-0.098

Dim_2	-0.1218	-0.986	0.037	0.1212	0.109
Dim_3	-0.0073	-0.05	0.631	-0.7271	-0.035
Dim_4	-0.0965	-0.056	0.679	0.5736	0.067
Dim_5	0.0357	0.003	0.155	0.1898	0.688

Correlation_of_Contributions_4000-word_sample_45_and_46

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.985	-0.042	0.28	0.284	-0.059
Dim_2	-0.066	0.987	0.18	0.093	0.135
Dim_3	0.388	-0.025	0.76	0.482	-0.012
Dim_4	0.162	0.355	0.36	0.699	0.065
Dim_5	-0.096	0.01	0.07	-0.011	0.859

Correlation_of_Coordinates_4000-word_sample_47_and_48

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.995	-0.164	-0.017	-0.098	0.083
Dim_2	-0.084	0.987	0.1	0.046	-0.037
Dim_3	0.012	-0.019	0.804	-0.378	-0.244
Dim_4	-0.095	0.089	0.192	0.892	-0.138
Dim_5	-0.078	0.074	0.339	0.283	0.244

Correlation_of_Contributions_4000-word_sample_47_and_48

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	-0.083	0.26	0.393	0.05436
Dim_2	-0.033	0.983	0.1	0.087	-0.00056
Dim_3	0.335	-0.015	0.79	0.536	0.17925
Dim_4	0.252	0.255	0.25	0.829	0.31145
Dim_5	-0.108	0.177	0.17	0.044	0.38501

Correlation_of_Coordinates_4000-word_sample_49_and_50

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.994	-0.1332	0.0857	-0.088	0.0671
Dim_2	-0.113	0.9874	-0.0167	0.047	-0.05
Dim_3	-0.08	0.121	-0.5889	0.737	0.0473
Dim_4	-0.002	0.0021	0.7504	0.487	0.0058
Dim_5	-0.013	-0.0079	-0.0066	0.094	0.5481

Correlation_of_Contributions_4000-word_sample_49_and_50

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	-0.086	0.35	0.161	0.023
Dim_2	-0.052	0.986	-0.05	0.254	0.22
Dim_3	0.345	0.245	0.54	0.62	0.158
Dim_4	0.368	-0.055	0.75	0.35	0.061
Dim_5	0.041	-0.05	0.1	0.012	0.365

Correlation_of_Coordinates_5000-word_sample_51_and_52

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.997	-0.116	0.03	-0.07	0.0848
Dim_2	-0.12	0.992	0.059	0.076	-0.0024
Dim_3	-0.106	0.103	-0.103	0.873	0.0448
Dim_4	0.017	-0.009	0.914	-0.179	0.0906
Dim_5	0.049	0.034	0.059	-0.066	0.9331

Correlation_of_Contributions_5000-word_sample_51_and_52

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.989	-0.067	0.3946	0.22	-0.03
Dim_2	-0.06	0.987	-0.0092	0.33	0.107
Dim_3	0.219	0.247	0.3637	0.9	0.195
Dim_4	0.386	-0.081	0.9629	0.3	0.021
Dim_5	-0.018	0.08	0.0029	0.2	0.939

Correlation_of_Coordinates_5000-word_sample_53_and_54

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9961	-0.1174	-0.058	0.096	0.054
Dim_2	-0.1246	0.9891	0.064	0.015	0.012
Dim_3	-0.0025	-0.005	0.412	0.79	-0.061
Dim_4	-0.1068	0.1332	0.773	-0.542	0.049
Dim_5	-0.0799	0.108	0.291	0.05	0.599

Correlation_of_Contributions_5000-word_sample_53_and_54

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.991	-0.061	0.206	0.4	-0.0204
Dim_2	-0.057	0.99	0.276	-0.032	0.0081
Dim_3	0.374	-0.035	0.386	0.78	-0.0053
Dim_4	0.224	0.277	0.667	0.527	0.2892
Dim_5	-0.029	0.067	0.084	0.028	0.7332

Correlation_of_Coordinates_5000-word_sample_55_and_56

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.996	-0.089	0.021	-0.106	0.04013
Dim_2	-0.137	0.99	0.077	0.088	-0.01878
Dim_3	0.046	-0.019	0.894	-0.18	-0.1783
Dim_4	-0.098	0.078	-0.025	0.942	0.00036
Dim_5	0.075	0.014	0.231	-0.018	0.79057

Correlation_of_Contributions_5000-word_sample_55_and_56

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.991	-0.0706	0.322	0.21	-0.014
Dim_2	-0.055	0.9933	-0.026	0.27	0.104
Dim_3	0.415	0.0016	0.958	0.48	0.085
Dim_4	0.308	0.1899	0.371	0.93	0.205
Dim_5	-0.017	0.1385	0.017	0.28	0.735

Correlation_of_Coordinates_5000-word_sample_57_and_58

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.996	-0.12	0.093	-0.01621	0.034
Dim_2	-0.122	0.984	-0.091	0.0006	-0.05
Dim_3	0.019	0.1	0.669	0.64983	0.046
Dim_4	-0.095	0.093	-0.707	0.59729	0.075
Dim_5	0.053	0.039	0.107	0.11162	0.918

Correlation_of_Contributions_5000-word_sample_57_and_58

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9916	-0.073	0.3417	0.15	-0.031
Dim_2	-0.0588	0.987	-0.0043	0.38	0.057

Dim_3	0.3524	-0.044	0.6785	0.42	0.102
Dim_4	0.2788	0.19	0.6221	0.52	0.285
Dim_5	-0.0059	0.101	0.0361	0.18	0.956

Correlation_of_Coordinates_5000-word_sample_59_and_60

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.996	-0.136	-0.085	0.02	0.043
Dim_2	-0.105	0.987	0.065	0.038	-0.014
Dim_3	-0.021	0.073	0.131	0.868	0.113
Dim_4	-0.098	0.108	0.933	-0.244	-0.036
Dim_5	0.111	-0.132	-0.207	-0.062	0.555

Correlation_of_Contributions_5000-word_sample_59_and_60

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	-0.066	0.33	0.322	-0.034
Dim_2	-0.061	0.991	0.14	0.02	0.101
Dim_3	0.235	0.164	0.36	0.898	0.109
Dim_4	0.329	0.081	0.92	0.285	0.115
Dim_5	0.151	0.289	0.28	0.036	0.214

Correlation_of_Coordinates_6000-word_sample_61_and_62

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.996	-0.108	-0.0045	-0.112	0.107
Dim_2	-0.135	0.989	0.1106	0.056	-0.016
Dim_3	-0.002	-0.021	0.8808	0.11	0.173
Dim_4	-0.089	0.119	-0.1962	0.899	-0.18
Dim_5	-0.016	0.051	-0.0662	0.293	0.475

Correlation_of_Contributions_6000-word_sample_61_and_62

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9929	-0.074	0.334	0.32	-0.0076
Dim_2	-0.0368	0.988	0.064	0.17	0.1841
Dim_3	0.2789	0.116	0.921	0.31	0.0164
Dim_4	0.3128	0.135	0.436	0.91	0.1827
Dim_5	-0.0077	0.036	0.026	0.11	0.5143

Correlation_of_Coordinates_6000-word_sample_63_and_64

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.995	0.151	-0.1	-0.019	-0.0059
Dim_2	-0.087	-0.988	0.12	0.028	0.0725
Dim_3	0.047	-0.041	-0.24	0.919	0.144
Dim_4	-0.066	-0.027	0.91	0.042	0.2674
Dim_5	0.089	0.049	-0.2	-0.041	0.8558

Correlation_of_Contributions_6000-word_sample_63_and_64

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	-0.033	0.34	0.3149	-0.0735
Dim_2	-0.08	0.983	0.22	-0.0021	0.0957
Dim_3	0.369	0.062	0.51	0.9509	-0.0077
Dim_4	0.256	0.231	0.89	0.3387	0.1697
Dim_5	-0.024	0.065	0.16	0.1403	0.9252

Correlation_of_Coordinates_6000-word_sample_65_and_66

	Dim_1	Dim_2	Dim_3	Dim_4Dim_5	
Dim_1	0.9969	-0.1606	0.012	-0.073	0.066
Dim_2	-0.0809	0.9898	0.034	0.063	0.011
Dim_3	-0.0251	0.0662	0.769	0.441	0.024
Dim_4	0.0998	-0.0694	0.526	-0.803	0.082
Dim_5	0.0026	0.0068	0.118	0.174	0.914

Correlation_of_Contributions_6000-word_sample_65_and_66

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.99	-0.064	0.368	0.32	-0.022
Dim_2	-0.061	0.988	-0.013	0.18	0.091
Dim_3	0.155	0.263	0.61	0.36	0.115
Dim_4	0.34	0.034	0.593	0.73	0.052
Dim_5	-0.082	0.1	0.043	0.15	0.921

Correlation_of_Coordinates_6000-word_sample_67_and_68

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.996	-0.103	-0.069	-0.0804	0.073
Dim_2	-0.137	0.988	0.058	0.0596	-0.022
Dim_3	-0.055	0.079	0.922	0.1415	0.166
Dim_4	0.068	-0.015	0.062	-0.9619	0.022
Dim_5	-0.018	0.076	0.1	0.0094	0.747

Correlation_of_Contributions_6000-word_sample_67_and_68

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.992	-0.063	0.183	0.3671	-0.042
Dim_2	-0.057	0.986	0.213	-0.0064	0.086
Dim_3	0.124	0.325	0.953	0.0046	0.137
Dim_4	0.358	-0.021	0.019	0.9728	0.031
Dim_5	-0.021	0.053	0.049	-0.0185	0.789

Table 54: Correlation Matrices of Dimension coordinates and contributions from MDAs of samples of Trolling tweets

Correlation_of_Coordinates_500-word_Troll_Sample_1_and_2

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.966	-0.0094	0.086	0.02	0.033
Dim_2	-0.023	-0.7176	0.272	-0.029	0.138
Dim_3	0.141	0.25	0.431	0.125	0.294
Dim_4	0.031	-0.0345	0.268	0.264	-0.062
Dim_5	0.045	0.088	-0.113	0.292	0.317

Correlation_of_Contributions_500-word_Troll_Sample_1_and_2

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.925	0.121	0.244	0.319	0.06
Dim_2	0.161	0.719	-0.01	0.18	-0.093

Dim_3	0.302	0.146	0.675	-0.054	0.129
Dim_4	0.015	0.139	0.217	0.2	0.238
Dim_5	0.148	0.093	0.244	0.117	0.352

Correlation_of_Coordinates_500-word_Troll_Sample_3_and_4

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9714	0.025	0.059	0.04	-0.01026
Dim_2	0.0951	-0.734	0.034	-0.088	-0.00012
Dim_3	0.0819	0.209	0.399	-0.639	-0.01595
Dim_4	0.0152	0.101	0.591	0.116	0.03432
Dim_5	-0.0077	0.082	0.132	0.024	0.42832

Correlation_of_Contributions_500-word_Troll_Sample_3_and_4

	Dim_1	Dim_2	Dim_3	Dim_4Dim_5	
Dim_1	0.916	0.19	0.129	0.18	0.016
Dim_2	0.363	0.59	0.03	0.26	0.019
Dim_3	0.068	0.32	0.103	0.62	0.435
Dim_4	0.224	0.25	0.664	0.17	0.131
Dim_5	-0.025	0.3	0.144	0.14	0.487

Correlation_of_Coordinates_500-word_Troll_Sample_5_and_6

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.971	0.107	0.13	0.05	0.0312
Dim_2	0.048	-0.646	0.103	-0.12	0.3158
Dim_3	0.022	0.374	0.537	-0.22	-0.001
Dim_4	0.086	-0.022	-0.068	0.14	-0.1952
Dim_5	0.044	0.067	0.2	0.1	-0.1014

Correlation_of_Contributions_500-word_Troll_Sample_5_and_6

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.917	0.15	0.197	0.251	-0.072
Dim_2	0.213	0.61	0.174	0.201	0.186
Dim_3	0.179	0.25	0.717	0.129	0.05
Dim_4	0.06	0.25	0.096	0.033	0.173
Dim_5	-0.109	0.11	0.027	-0.019	0.297

Correlation_of_Coordinates_500-word_Troll_Sample_7_and_8

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.977	-0.0513	0.092	-0.027	-0.042
Dim_2	0.08	-0.6566	0.115	0.076	-0.132
Dim_3	0.146	-0.0082	0.357	-0.196	0.413
Dim_4	-0.04	-0.0864	0.208	0.229	0.19
Dim_5	0.038	-0.2534	-0.086	0.336	0.036

Correlation_of_Contributions_500-word_Troll_Sample_7_and_8

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.922	0.172	0.3	-0.055	0.014
Dim_2	0.124	0.623	0.2	0.194	0.024
Dim_3	0.301	0.173	0.37	0.271	0.163
Dim_4	0.033	-0.096	0.11	0.211	0.147
Dim_5	-0.017	0.265	0.26	0.26	0.147

Correlation_of_Coordinates_500-word_Troll_Sample_9_and_10

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9817	0.068	0.089	-0.0001	0.085
Dim_2	0.0021	0.765	0.166	0.0243	-0.062
Dim_3	-0.0072	-0.215	0.221	0.37828	0.021
Dim_4	-0.022	-0.051	0.263	0.0499	-0.329
Dim_5	0.0371	0.171	0.08	-0.09838	-0.176

Correlation_of_Contributions_500-word_Troll_Sample_9_and_10

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9438	0.25	0.2144	0.12	-0.021
Dim_2	0.0075	0.647	0.133	0.17	0.059
Dim_3	0.1665	0.115	0.3443	0.52	0.209
Dim_4	0.1603	-0.046	0.2707	0.14	0.085
Dim_5	0.092	0.259	0.0029	0.28	0.413

Correlation_of_Coordinates_1000-word_Troll_Sample_11_and_12

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.984	-0.009	0.093	0.074	0.092
Dim_2	0.075	0.855	0.19	-0.08	0.182
Dim_3	0.056	-0.11	0.684	-0.176	0.21
Dim_4	0.029	-0.117	0.086	0.369	-0.252
Dim_5	-0.064	-0.102	0.224	-0.178	-0.413

Correlation_of_Contributions_1000-word_Troll_Sample_11_and_12

	Dim_1	Dim_2	Dim_3	Dim_4Dim_5	
Dim_1	0.93	0.2044	0.167	0.081	0.069
Dim_2	0.2	0.8179	0.089	0.162	0.26
Dim_3	0.273	0.0082	0.782	-0.069	0.23
Dim_4	0.29	0.2024	0.022	0.222	0.352
Dim_5	0.033	0.1109	0.129	0.084	0.325

Correlation_of_Coordinates_1000-word_Troll_Sample_13_and_14

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.988	0.045	0.071	-0.019	-0.068
Dim_2	-0.011	0.86	0.03	-0.11	-0.043
Dim_3	0.032	-0.069	0.762	-0.163	0.039
Dim_4	-0.02	-0.162	0.019	0.155	-0.083
Dim_5	-0.035	-0.047	0.123	0.09	0.308

Correlation_of_Contributions_1000-word_Troll_Sample_13_and_14

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.968	0.322	0.264	0.15	0.033
Dim_2	0.213	0.751	0.108	0.52	-0.042
Dim_3	0.248	0.124	0.863	0.12	0.02
Dim_4	0.067	0.249	0.041	0.1	0.246
Dim_5	0.063	0.092	0.062	0.21	0.081

Correlation_of_Coordinates_1000-word_Troll_Sample_15_and_16

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.987	0.024	0.11	0.019	-0.066
Dim_2	0.027	0.844	-0.26	-0.1144	0.081
Dim_3	0.019	0.172	0.52	-0.0415	0.42

Dim_4	0.115	-0.049	0.31	0.1706	-0.493
Dim_5	-0.069	-0.253	-0.22	0.0077	0.209

Correlation_of_Contributions_1000-word_Troll_Sample_15_and_16

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9739	0.157	0.21	0.112	-0.14
Dim_2	0.2924	0.792	0.13	0.402	0.01
Dim_3	0.2396	-0.044	0.87	0.068	0.13
Dim_4	0.2251	0.093	0.12	0.044	0.3
Dim_5	-0.0093	0.294	0.17	0.076	0.33

Correlation_of_Coordinates_1000-word_Troll_Sample_17_and_18

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9861	0.00082	0.06	0.045	0.0021
Dim_2	-0.0067	0.824	-0.349	-0.074	0.1014
Dim_3	0.0529	0.22065	0.522	-0.088	-0.195
Dim_4	0.1402	-0.09563	0.151	0.585	0.2956
Dim_5	-0.004	-0.09255	0.085	-0.381	0.3382

Correlation_of_Contributions_1000-word_Troll_Sample_17_and_18

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.94	0.103	0.34	0.064	0.2146
Dim_2	0.24	0.762	0.225	0.185	0.2343
Dim_3	0.2	0.027	0.707	0.178	0.008
Dim_4	0.27	0.225	0.015	0.471	0.1
Dim_5	0.16	0.28	0.036	0.38	0.1492

Correlation_of_Coordinates_1000-word_Troll_Sample_19_and_20

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9886	-0.017	0.078	0.03	-0.0588
Dim_2	0.0764	0.871	-0.077	-0.158	-0.2117
Dim_3	0.1462	0.266	0.69	-0.282	0.0085
Dim_4	0.0028	0.051	0.111	0.514	0.3908
Dim_5	0.0943	-0.06	0.217	0.182	-0.3215

Correlation_of_Contributions_1000-word_Troll_Sample_19_and_20

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.96	0.294	0.2221	0.048	-0.13
Dim_2	0.23	0.783	0.2189	0.257	0.201
Dim_3	0.19	0.095	0.8489	0.03	0.107
Dim_4	0.29	0.293	0.1165	0.352	0.018
Dim_5	0.14	0.074	0.0084	0.258	0.334

Correlation_of_Coordinates_2000-word_Troll_Sample_21_and_22

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9914	-0.0003	0.071	0.0029	-0.038
Dim_2	0.0154	0.8845	-0.116	-0.0509	0.142
Dim_3	0.138	0.1455	0.786	0.1305	-0.154
Dim_4	0.002	-0.2174	-0.271	0.5536	-0.419
Dim_5	0.0936	0.1374	-0.178	-0.4667	-0.345

Correlation_of_Contributions_2000-word_Troll_Sample_21_and_22

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5

Dim_1	0.97	0.222	0.22	0.061	-0.037
Dim_2	0.31	0.854	0.2	0.314	0.066
Dim_3	0.32	0.106	0.86	0.109	-0.01
Dim_4	0.13	0.412	0.16	0.501	0.638
Dim_5	0.12	0.099	0.14	0.36	0.401

Correlation_of_Coordinates_2000-word_Troll_Sample_23_and_24

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.99	0.0059	0.106	0.019	-0.04
Dim_2	0.025	0.919	0.309	-0.103	-0.041
Dim_3	0.079	-0.2407	0.798	0.045	0.051
Dim_4	-0.025	-0.2093	-0.146	0.78	0.13
Dim_5	0.102	-0.1321	0.072	0.151	-0.438

Correlation_of_Contributions_2000-word_Troll_Sample_23_and_24

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.98	0.13	0.25	0.2096	-0.077
Dim_2	0.355	0.77	0.09	0.4748	0.021
Dim_3	0.228	0.15	0.93	0.001	0.099
Dim_4	0.031	0.5	0	0.6509	0.178
Dim_5	0.138	0.13	0.11	0.2495	0.43

Correlation_of_Coordinates_2000-word_Troll_Sample_25_and_26

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.995	0.038	0.101	0.0158	0.0155
Dim_2	-0.0084	0.926	-0.06	-0.0098	-0.0088
Dim_3	0.1029	0.137	0.784	-0.0913	-0.2831
Dim_4	0.0656	-0.226	-0.214	0.4041	-0.3769
Dim_5	-0.0636	-0.144	0.297	0.3715	0.3503

Correlation_of_Contributions_2000-word_Troll_Sample_25_and_26

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.979	0.204	0.19	0.174	0.059
Dim_2	0.259	0.892	0.112	0.313	0.131
Dim_3	0.325	0.089	0.899	-0.015	-0.009
Dim_4	0.076	0.23	0.097	0.49	0.62
Dim_5	-0.024	0.168	0.106	0.344	0.385

Correlation_of_Coordinates_2000-word_Troll_Sample_27_and_28

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.9925	0.03	0.1	0.00046	-0.029
Dim_2	0.0082	0.915	-0.28	-0.15858	0.083
Dim_3	0.1109	0.29	0.78	-0.18612	0.037
Dim_4	0.0448	0.014	0.15	0.66898	-0.046
Dim_5	0.0997	0.088	-0.15	-0.10814	-0.411

Correlation_of_Contributions_2000-word_Troll_Sample_27_and_28

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.981	0.276	0.2652	0.0918	-0.122
Dim_2	0.218	0.904	0.148	0.4002	0.01
Dim_3	0.233	0.038	0.9255	0.0052	0.01
Dim_4	0.232	0.454	-0.0096	0.6957	0.075

Dim_5	-0.034	-0.024	0.0957	0.112	0.666

Correlation_of_Coordinates_2000-word_Troll_Sample_29_and_30

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.994	-0.0017	0.0728	0.0024	-0.11
Dim_2	0.016	0.9216	0.1277	-0.1697	-0.16
Dim_3	0.11	-0.0715	0.7911	-0.0907	-0.03
Dim_4	0.047	-0.1477	0.0834	0.7067	0.34
Dim_5	-0.068	0.1237	0.0036	-0.09	0.48

Correlation_of_Contributions_2000-word_Troll_Sample_29_and_30

	Dim_1	Dim_2	Dim_3	Dim_4	Dim_5
Dim_1	0.98	0.275	0.2833	0.188	-0.0061
Dim_2	0.25	0.924	0.0773	0.499	0.0683
Dim_3	0.22	0.119	0.9171	-0.066	0.0016
Dim_4	0.13	0.384	-0.0019	0.71	0.2217
Dim_5	-0.13	-0.014	0.0798	0.09	0.5358

Appendix 3: Extra Tweets Associated with the General Twitter's Dimensions of Linguistic Variation

POSITIVE: Dimension 2	coord	ctr
· LISTEN: John Hodgson discusses "the emperor of all conjurers" RICHARD POTTER on @nhpr: https://t.co/uJPyaiwhYy (37 minutes in). Read more about this book which the @WSJ called a "provocative record of Potterís odds-defying climb" here: https://t.co/zoAxS24sfK #magic #celebrity	0.811	0.099
· farmdoc Webinar: @jt_hubbs and @ScottIrwinUI will review @USDAís June 29 Grain Stocks and Acreage reports and considers balance sheet and price implications for both old and new crop #corn and #soybeans. Register here: https://t.co/pqovl299qH https://t.co/dLtJZRM5HY	0.802	0.097
· VIDEO - @odublast - "Building Leaders for Advancing Science and Technology." High school students from across the Commonwealth spent 3 days learning about to climate change, sea level rise and cybersecurity through STEM. WATCH NOW: https://t.co/YdaYTB7fJw #odu @educationODU https://t.co/DfCATdltVd	0.757	0.086
· The Phoenix @Suns had a huge profile in yesterdayís #NBADraft2018 with four picks, a big trade and brand new big man in @DeandreAyton with roots here in #Arizona. Details ahead @kjzzphoenix. Stream us here: https://t.co/qd2LqqGXs4.	0.756	0.086
· Just signed up for WCX, the global digital currency exchange. Sign up & earn 100 X Tokens: https://t.co/83d2f2xOZj @wcxofficial #TradeTheWorld	0.747	0.084
· Next Friday at @ICALondon, catch 'Daughters of Africa Screen Narratives: Archive Revelations VII': an event curated by June Givanni in support of @JG_PACA featuring ten short films made by African	0.746	0.084

women. See the full programme here: https://t.co/SX1UuFTybh		
· Welcome to one of our newest #FabFem role models, Alexandra Forsythe! Alexandra is currently working three jobs: (1) electrical engineering contractor for Ultra Electronics (USSI) designing, building, and testing electronics;... https://t.co/mmBHDHUWqc	0.743	0.083
· UP Next:11:15 a.m.-12:30 p.m. Workshops From #metoo to Meaningful Change: Sexual Harassment in the Workplace - Room 007Information Privacy & Security Checklists for Startups - Room 264The Scrappy, People-Centered Startup - Room 209	0.737	0.081
· In Chicago News: The fast pace of global economic change, and in the logistics supporting economic activity, requires equal levels of responsiveness in real estate. Connect Industrial, July 11 in Chicago, will focus on this issue. #CRE #ConnectIndustrial https://t.co/lnsTchaMHx	0.736	0.081
· Baldur's Gate II: Throne of Bhaal debuted 17 years ago! Enjoy the final chapter in the #BaldursGate story as a part of Baldur's Gate II: Enhanced Edition! Dare you claim your fatherís throne?https://t.co/G7IdTlqMc5 https://t.co/fmbEpYI4Pq	0.7	0.074
· RT @kgauravITC: Part 2 of @Investingcom's recent article on how @CashaaLtd is solving the problems of the #unbanked population: https://t.cÖ https://t.co/2MRIaqbE8W	0.699	0.073
· Host nation Russia scored five goals against Saudi Arabia and another three against Egypt at this yearís #WorldCup. Hereís how the lowest ranked team became one of the biggest surprises so far in the tournament: https://t.co/G5zD6j1HSe https://t.co/lu1npWbsOe #mrs_bitcoin #newÖ	0.697	0.073
· Four buildings? Fifteen doors? One sad can of shaving cream? You thought #LukeCage needed a clean-up crew in the first season, just wait 'till the next one. See "Marvel's @LukeCage" destruction by the numbers so far: https://t.co/VJceNVvyCq https://t.co/VyqU9LBwxF	0.693	0.072
· Neymar is one the best....divers in the world. [emoji]#BRACOS #BRAxCRC #WorldCup: https://t.co/dsdxGw9vwU via @YouTube	0.691	0.072
· TAKE ACTION: Help fight the deadly wildlife #zombie disease! Tell Congress to pass legislation that will help #moose #deer #elk via @WildlifeAction https://t.co/7HGWfebU6N	0.683	0.07
· The#Noblesville4th Planning Committee has a limited number of 2018 collectible festival buttons ñ 500 2-inch, LED blinking buttons ($3) & 500 2.25-in round metal buttons ($1). Net proceeds from button	0.68	0.069

sales go directly to supporting July 4th festivities. https://t.co/F8IyRMcd0p https://t.co/fkfW5Og1lo		
· Tough words on UN #HumanRightsCouncil: The European Union should follow the U.S. decision and leave ìthat farce of a U.N. body called Human Rights Council,î said #Czech Member of the European Parliament @TomZdechovskyEP. @unhrcpr https://t.co/re7hHYXxfD	0.675	0.068
· Register for the IACC Regional Roundtable + gain access to FREE training with Stacey Geyer @MCA_training. Click here for details and to register: https://t.co/WUJNBrZVYN #eventprofs #meetingprofs https://t.co/WLfdMzxcPK	0.672	0.068
· [Latest Blog Post] SYNTASA NAMED A GARTNER COOL VENDOR IN PERSONALIZATION | Gartner has praised Syntasaís capabilities to combine clickstream + enterprise data to deliver a more holistic view of customer behavior: https://t.co/Bt6BKk5Nez | #Gartner #Personalization #AI https://t.co/UDp8JEsQ32	0.663	0.066
· RT @ShropCouncil"#FridayPhoto. Mowing hay in the Long Meadow, Brewers Oak, Shifnal. 1930. @ShropArchives ref: PH/S/14/1/36 https://t.co/fFD2D9AQlO"	0.659	0.065
· Now hiring for 273 #job opportunities at CVS Health (@CVSHealthJobs), LEGO Group (@LEGO_Careers), Acosta (@AcostaJobs), and more. https://t.co/99TlrSeOi1	0.658	0.065
· Happening tomorrow! Enjoy the wildflowers on a hike with @taylorheadpark on Sat., Jun 23 10am - 2pm. The wildflowers are in full bloom and the vegetation is fresh and lush. This is a walk everyone can enjoy, even those using mobility aids. https://t.co/tdGYHUGBYk @HikeNS #hikens https://t.co/G5t9zgPFx1	0.657	0.065
· ìDenial of reproductive rights is an issue of inequalityî ñ @Atayeshe As the threat of death during birth looms large for millions around the globe, our Executive Director Dr. Natalia Kanem says we must protect #reprorights for all: https://t.co/QJn0G6HxB2 @MediaplanetUSA #SRHR	0.657	0.065
· Oracle ERP Cloud helps @TonysChocoUS delivering on their mission: to produce and sell 100-percent slave-free chocolate. See how on this video. @Oracle @OracleERPCloud #OracleEmp https://t.co/8HxQ4nsIL5 https://t.co/lEbeGGrAS5	0.657	0.065
· #DidYouKnow that Italy has £5.1bn of investment in the UK? Trade Minister @GrahamStuart is in Milan promoting future investment opportunities for Italian companies in the UK including further	0.654	0.064

investment in the automotive and aerospace sectors. #FreeTradeUK https://t.co/eHlwq2Pwp8		
· A SURE HOUSEFreedomCourt at 10 AnniversaryCome and rejoice with us this SundayJune 24, 8:00amDresscode: ThanksgivingMinistering: Pastor Taiwo @taiwobolodeoku Bolodeoku, Pastor in Charge of RCCG JoshuaVille@thesegunaluko@CDistinguished @femi_Adejobi @liveway @myjoshuaville https://t.co/nvCi3CNsy8	0.652	0.064
· 12 Point Distillery, #LafayetteCOís new kid on the block, is featured in @5280Magazine this month. Itís a great weekend to explore their delicious concoctions and amazing outdoor patio! https://t.co/vOh0xBQwqk https://t.co/3vm6Rbf44q	0.649	0.063
· WIN a (4) pack of tickets to Magic Mountain in the 9a hour! 877-440-1047 #JimmyReyesInTheMorning @SixFlags @oldschool1047 https://t.co/cRrYOX9LK3	0.649	0.063
· The Healthwatch Portsmouth AGM takes place on Wednesday 27 June 6-8pm - we will be joined by the Director of @HealthwatchE and we will be celebrating #NHS70 with a cake! All welcome. See agenda + book place at https://t.co/gjDzlthjNN	0.647	0.063
· #Travel #Beach: VIANELLAS: Chic Flip-Flops to Wear on the Beach!... by @antondiaz, +6 more. https://t.co/uIRG0CzJUa #tbex	0.647	0.063
· #RVA https://t.co/OHKQKhLXUr Richmond Mag investigated the Little Saint controversy Business, Food, News, richmond, rva, va In 2016 the Tampa Bay Times featured a six-part series, ì Farm to Fable î exploring the local food movement, how restaurants jumped on board, and frequentÖ	0.64	0.061
· STOP FOCUSING ON Idaho or Texas!!!!! Focus on the rest of the clip where @foxandfriends and @kilmeade accuse the trump administration of illegal entry to mexico to kidnap a kid!!!!!! One is typical fox news the crime committed by the president's team. https://t.co/vftqlO4wIo	0.636	0.061
· MOS has the best action sequences ever made in any movie. Personal opinion best movie of all time. Zack+DC=EPIC @wbpictures @WarnerBrosEnt @DCComics @ATT bring Zack back and #ReleaseTheSnyderCut and continue on original plan https://t.co/DyXvHOdI8W	0.631	0.06
· Live: Durham Cricket Board Under 12 54/1 - 195/8 Yorkshire Cricket Board Under 12 B https://t.co/Asj7h4xXxd (via @ECB_cricket) #playcricket	0.631	0.06

· Tackling social problems through artistic intervention: @boukjecnossen studies the role of artefacts in the communicative constitution of the collective behind the ëCamping Kafkaí project. #OAP2018 @WorkshopOAP https://t.co/hrU1YMiljp	0.626	0.059
· For Sale: A 4 bedroom modern family home in Piggott Place, Sheet, #Petersfield| £675,000Viewings are highly recommended. Please call us now on 01730 233333 to arrange a viewing! #PropertyMore details at https://t.co/O1SjrBcPOD https://t.co/CvRcbDuUil	0.625	0.059
· Flood Advisory issued June 22 at 9:21AM CDT until June 23 at 1:00PM CDT by NWS: The National Weather Service in Chicago has issued a * Urban and Small Stream Flood Advisory for... De Kalb County in north central Illinois... Kane County in northeasternÖ https://t.co/fyTWLdCpbX	0.623	0.058
· Thank you @CareTalkMag for featuring #MyGPandMe. 9 MPs have so far signed the Early Day Motion (1365) to make learning disability training mandatory for student GPs. You can help. Here's how to speak to your MP to ask them to sign it: https://t.co/74Sl5ULmaD https://t.co/91e80Im35H	0.618	0.057
· In sharing the same #global vision, @UEuropea and @Fun_Realmadrid have ratified their collaboration agreement to sustain the Foundationís Social and Sports Schools that guarantee equal access to quality basketball training for over 500 boys and girls in Madrid. https://t.co/oJMMTHxU5x	0.618	0.057
· The migrant crisis signals an end to one era of Republicanismóand the terrifying start of a new one, writes @AlexWagner: https://t.co/jQjBy2Dsti	0.618	0.057
· Get building in your home environment with Architectural Lego! Sign up to the Synergy newsletter to win at the @VisionLDN ó stand V119. #VisionLDN #moretimefordesign https://t.co/UZTusDWGR7	0.615	0.057
· #ETHBuy at #Bitstamp and sell at #BTCTurk. Ratio: 2.11%Buy at #Bitstamp and sell at #Koineks. Ratio: 2.46%Buy at #Bitstamp and sell at #HitBTC. Ratio: 1.32%Buy at #Bittrex and sell at #HitBTC. Ratio: 0.92%#bitcoin #arbitrage #arbitraj #arbingtool https://t.co/xiFUPzcOcC	0.615	0.057
· Wikileaks has published professional information and LinkedIn profiles of thousands of US Immigration and Customs Enforcement employees in a searchable online database, as the fallout from the Trump administration's zero-tolerance policy continues. https://t.co/pHQ2aAfhWX https://t.co/RD4HLYTabx	0.612	0.056
· #OnTheGoVP-Acad organizes student fora The Vice President for Academics spearheaded a two-day student orientation held at the Center for Performing	0.612	0.056

Arts (CfPA) last June 19-20.Students were acquainted with the... https://t.co/kVwjAxoEFh		
· Closed a SELL USD/JPY position at 109.898 on ZuluTrade.PnL: 484.47USD Visit https://t.co/yp0lSMxpws to see my hypothetical performance.	0.612	0.056
· Closed a SELL USD/JPY position at 109.895 on ZuluTrade.PnL: 822.18USD Visit https://t.co/bFuJWCkx0u to see my hypothetical performance.	0.612	0.056
· Closed a SELL USD/JPY position at 109.895 on ZuluTrade.PnL: 516.34USD Visit https://t.co/bFuJWCkx0u to see my hypothetical performance.	0.612	0.056
· 180622 Jennieís Instagram (jennierubyjane): [emoji] [emoji] [emoji]ïJennieís latest post reached 1M likes in just 4 hours!!! The fastest!!! [emoji] [emoji] [emoji] [emoji] [emoji] [emoji]#BLACKPINK #JENNIE https://t.co/CTSOSU8bOW	0.61	0.056
· 15:30 #Fairview1st 10 Reine Tonnerre 11/12nd 4 Excellent 22/13rd 13 Maverick Girl 11/116 Ran.SF: 249.23†ZAR [emoji] https://t.co/t5HCJXWccN [emoji] https://t.co/29uFfpboCg [emoji] https://t.co/P0cmg0oyCohttps://t.co/t5HCJXWccN	0.608	0.056
· Last year, two New York real estate developers, Kenneth Nakdimen and Shalom Lamm, pleaded guilty to interfering with a mayoral election in Bloomingburg, New York, including using fake...Contact me for more info: Everette Upsher ï (757) 848-8011 ï... https://t.co/rebmwZLvtz	0.607	0.055
· ìMina ngifuna umtwana oYello ... being unruly at @Cannes_Lions - after that level of creativity one needs to chill now and enjoy the best of the French Riviera ... [emoji] Nabo babes abaNice ...î - #LIONCanAtCannesLION https://t.co/4dDkySKDXU	0.606	0.055
· Lee & Associates Arizona reports the Phoenix #industrial market continues its positive momentum after a record year in 2017, with 1,359,837 SF of positive absorption in Q1 2018. https://t.co/bJA4dEB264	0.604	0.055
· ATTI: #MLBG to Life Sc & Towers Bus 292 is at Grant@4th at 6/22/2018 11:08:58 AM. Next Stop: Grant@1st.	0.602	0.054
· [emoji] [emoji] United Kingdom [emoji] ASCOT (Race 2) [emoji] 21/06/2018 [emoji] 14:05 GMT 1st HUNTING HORN (4)2nd CROSSED BATON (1)3rd ZAAKI (16)Congratulations to Ryan Moore (Jockey) and All Connections! #Ascot#HorseRacing#Results https://t.co/T9QnpLEXyR	0.602	0.054

· A great read to push you through your Friday afternoon into the WEEKEND!50 Innovation and Success Quotes from SpaceX Founder Elon Musk > >https://t.co/xYC616nlH6 Ö #Tesla #innovation #FridayFeeling #FridayReads #motivational #success #weekend https://t.co/ZKaPwR0qS1	0.601	0.054
· #NowPlaying Bow 'N' Arrow [FINAL] [JPA] - Robb Mykes AKA Masked Reality On Go Global Radio For AirPlay email: Goglobalradio@gmail.com for more info@Goglobalradio	0.601	0.054
· M. LOUISE FITZPATRICK, EDD, RN, FAAN (1942 -2017) - HALL OF FAME AWARD RECOGNIZES THE LIFELONG COMMITMENT OF INDIVIDUAL NURSES TO THE PROFESSION OF NURSING AND THEIR IMPACT ON THE HEALTH & THE SOCIAL HISTORY OF THE U.S.#MA2018 #NURSE https://t.co/IQa1QJrzMK	0.6	0.054
· Welcome home, White Team! Since 2006, the @1stTSC has been perpetually deployed to the @CENTCOM region, making sure that warfighters have the supplies & capabilities needed to accomplish their missions. Every 6 months, they deploy their Red, White, or Blue team to the region. https://t.co/OENTF3FzS9	0.597	0.053
· Check out what I just added to my closet on Poshmark: Vaneli Calf Hair & Leather Flats - Size 8 1/2M. https://t.co/Wxtohls33s via @poshmarkapp #shopmycloset	0.597	0.053
· In @TheHill, Dr. Vinita Parkash writes about the epidemic of healthcare worker suicide. The high incidence of depression and #PhysicianBurnout are culprits. MDsyncNETís communications modules increase physician satisfaction. https://t.co/vWumz7m3Tm https://t.co/VNJuzwVocg	0.597	0.054
· NEW -#RanchFire #NM #NMGNF https://t.co/p7BWaa7WKJ Ranch Fire: The Ranch Fire was reported at approximately 3:15 p.m. on June 21\, and appear	0.596	0.053
· 2018 #NBA Draft: 5 first round steals https://t.co/yjoKKo9vCf -- @MaxSHHolm via @HoopsHabit https://t.co/GjkAwSmtgR	0.596	0.053
· A true showcase of the strength in partnership between GCI & leading customer communications experts, @EnghouseInterac, @DimensionsUK's #ContactCentre solution not only affords the company greater cost savings, but also offers an improved overall interactive customer experience! https://t.co/bTVZCkPePy	0.594	0.053
· See our latest England #job and click to apply: Barista - Store# 12170, CANTERBURY-SNSBRY KN -	0.594	0.053

https://t.co/mKDswurNS3 #CustomerService #Hiring #CareerArc		
· The Bridge to the #Digital World - The @sensorplustest from 26.06. to 28.06. in Nuremberg, today, we invite you to the leading forum for sensors measuring and testing technologies. Join the #Innovation Dialog with your free ticket! https://t.co/H8f7p4rMYY #sponsored_ad #vfv18 https://t.co/b9xh8JbpsG	0.593	0.053
· Ridiculously light. Seriously thin. Win An Apple iPad Mini 4 (Worth $399) thanks to @PrizeTopia. Enter here: https://t.co/sXOHbtpDQq Ö #PrizeTopia #Apple #iPadmini #ipadmini4 #iPad #giveaway #prize #competition #contest #freestuff @Apple @AppleMusic @AppleNews	0.593	0.053
· Ridiculously light. Seriously thin. Win An Apple iPad Mini 4 (Worth $399) thanks to @PrizeTopia. Enter here: https://t.co/2Fk9GtXcWV Ö #PrizeTopia #Apple #iPadmini #ipadmini4 #iPad #giveaway #prize #competition #contest #freestuff @Apple @AppleMusic @AppleNews	0.593	0.053
· [emoji]THE SHADES [emoji] : NOW OPEN FOR AUDITION : 18-23 JUNE 2018 FORM >> https://t.co/PBDYJD6kBW#KRISHADES #KAISHADES #HUNSHADES # [emoji] [emoji]forsex # [emoji] [emoji]forhost # [emoji] [emoji] [emoji] [emoji] # [emoji]img # [emoji] [emoji] https://t.co/XI9dYO4k7u	0.593	0.053
· Now: @SangerNYT joins to talk about America's vulnerability to cyberattacks by global adversaries, covered in his new book "The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age." Tune in at https://t.co/JE6GnVgu5p.	0.592	0.053
· [emoji] Happn takes on Tinder Places with an interactive map of missed connectionshttps://t.co/rBoWyJPUFw #apps Article Published on June 21, 2018@2:08pm	0.591	0.053
· Soundzrise LIVE Sunday 24th June 2018 The Ritual with AnanË & Louie Vegafrom 6pm to 00 @ Riva Beach Club Fregene powered by: ONSET - KNM Music - Loud Professional - ElectroVinyl - @DEEPPOSO Æ - Anarchy in the Club - Soundzrise https://t.co/cYn7izuyT0	0.589	0.052
· It's just over a month until #2018AACC officially open it's doors! Visit the @VitlProducts team at BOOTH 1876 to find out more about our latest products including the Lu-mini & our extended range of Heated Modules: https://t.co/wll28PWK5k @_AACC #lifesciences #labware https://t.co/eIZ78Zuh0b	0.588	0.052

· MLB Alert 6/22/18 11: Looking for a new daily fantasy site. Try DRAFT: https://t.co/V6Wfz9N4CQ	0.587	0.052
· ANALYSIS: As countdown to Paris protest begins, is ëFree Iraní the only alternative https://t.co/k607S3TTYO via @AlArabiya_Eng#FreeIran2018 #IranRegimeChange	0.585	0.051
· BET NOW with #Matchbook and get GBP500 cash back in your first 5 weeks. -18+,Ts&Cs apply-https://t.co/qHRT72R7BQ https://t.co/jpQFLmsvUn	0.584	0.051
· So @wave5trade Called SHORT $LEN on 18 April - Smashed Through Target 19 April! Get #Stocks #Trading Journal including link to original Signals Video >>HERE>> https://t.co/tnIXT7eS04 #tradethe5th #NinjaTrader #Thinkorswim #TradeStation #MultiCharts https://t.co/dtDmEqVMd9	0.58	0.05
· [emoji] [emoji] France [emoji] LA TESTE-DE-BUCH (Race 1) [emoji] 21/06/2018 [emoji] 10:40 GMT 1st CNICHT (5)2nd GAILLEFONTAINE (7)3rd BHARUCH (4)Congratulations to Roberto .C Montenegro (Jockey) and All Connections! #LA TESTE-DE-BUCH#HorseRacing#Results https://t.co/7YZ1uzSbzK	0.58	0.05
· Reputed PR and Advertising Agency required #KU Mass Communication fresh graduates. Kindly send your request via inbox. #MCD Job Role: P.R. Executive. (Client Services - Corporate)Job Description/Responsibility... https://t.co/ATJ6CEPO4W	0.579	0.05
· [emoji]WC19 now available on the Edge 3! [emoji] A WC19 crash-test approved occupied transit option is now available on the Edge 3 with captainís seating, SynergyÆ seating, and TRU-BalanceÆ 3 seating configurations. https://t.co/L0cO9sXrOi	0.577	0.05
· What if the cast of #LordOfTheFlies were all women? An interesting exam question? Book your school in now to see our co-production with @ShermanTheatre and get a FREE interactive workshop! https://t.co/aIEDNBCYIX #englishgcse @wjec_cbac @CCEA @Edexcel https://t.co/FtuEQNFHgj	0.576	0.05
· Modest Dept on location in Berlin for a Puma shoot with their Hypercore powered @smallhd & @teradek 703 Bolt, @RED_Cinema Scarlet-W, and @freeflysystems MoVI Pro. https://t.co/DsuvipZS3U	0.576	0.05
· Congrats to BenU Alum, Michelle Allen, who was chosen this week among the "40 Under 40 Emerging Nurse Leaders" by the #INF. Michelle earned her MS in Nursing degree at BenU in 2014 and has worked as both an RN and an adjunct professor. [emoji] [emoji] https://t.co/4kEHHrTnr2Ö/40-under-40/	0.575	0.05

· Information letter on ActilyseÆ powder and solvent for solution for injection and infusion (alteplase) for acute ischaemic stroke: Important extension of indication to include adolescents under 16 years of age https://t.co/dm03Ib5Sk0 #pharmacovigilance	0.574	0.049
· #BRA vs #CRC: Brazil beat #CostaRica 2-0 in tense encounter https://t.co/XQxfoSOJGl #WorldCup #soccer #football #sports @TOISports	0.574	0.049
· If you want to support an awesome young person running for #Congress, CLICK HERE to support @garethtrhodes, endorsed by #NYTimes and thousands of people that chipped in $19 each (he probably needs some bigger checks now to get the msg out) https://t.co/RzXxVEnxwA	0.572	0.049
· Attending #ISTE18? Here are 5 places #teachers and education #entrepreneurs should visit in #Chicago: https://t.co/ApDtH7ctET	0.57	0.049
· Regina SK Weather, Temp:22.2∫C; Dew:14.3∫C; Pressure:1004.70hPa.; Wind:0/kph@80.0; Humdity:61 https://t.co/6Ptyl5RMRM https://t.co/vfZcUnWHSI	0.57	0.049
· Regina SK Weather, Temp:22.5∫C; Dew:14.9∫C; Pressure:1004.60hPa.; Wind:0/kph@80.0 Wind Chill:22.4∞C Humidity:62; https://t.co/6Ptyl5RMRM https://t.co/vfZcUnWHSI	0.57	0.049
· $TSLA Tesla's map of their charging infrastructure that are built in North America, about 600 stations and 5,000 chargers. (The numbers below, 1,261 and 10,021, are global figures. Red pins are operational today, grey pins should be operational within about a year). https://t.co/HIJ8avXfnu	0.569	0.049
· 3 of the last remaining Crooners heading down to Wimborne for some serious Croonin No G with our 9-Piece Big Band #seriousjockinAt the @TivoliWimborne ! Pip pip! https://t.co/1nfjJJWY0h	0.569	0.049
· Win Curse of the Ancients by #HawkMacKinney + $25 Amazon GC @iReadBookTours #CurseoftheAncients https://t.co/9BlRlaE8su	0.569	0.049
· The latest The StemCELLS 21 Daily! https://t.co/Q0hEMsniwu Thanks to @QSTwits @DriverBrian #news #bbcworldcup	0.569	0.049
· The level of creativity that has come out from these pitchers at the @sic_NG today is on a whole different level.. I am truly impressed, and my hope in the future of Nigeria is renewed.#SICNig#SouthWestSicNG https://t.co/sDvqd6Yn0A	0.567	0.048
· The DreamPort #website is now live! DreamPort is a combination of state-of-the-art facilities, #innovative programs, and imaginative people charged with finding that spark that leads to unparalleled	0.567	0.048

capability for @USCYBERCOM and the #warfighters at large. https://t.co/gSr1q6tXdK https://t.co/2bhxYk7lZ2		
· #FridayFeeling: THANKFUL for all of the marketing industry experts who endorsed my book. 100% of my author proceeds are donated to charities blessing children in need. https://t.co/Vl7VpWwrqs #Marketing #Advertising #ClientRelationships #Proactivity #Training #MarketingConsultant https://t.co/GN85VDwfm2	0.567	0.048
· Speakers at press conference for #ProjectPatton Friday June 22, 2018 ~ Toronto Police Deputy Chief James Ramer and Acting Inspector Donald Belanger of the Integrated Gun & Gang Task Force. Learn more about gang prevention, intervention and suppression: https://t.co/WyBpTbpWJ8 ^sm https://t.co/TCg3MqmxIW	0.566	0.048
· Announcing Regional Winner - Amanda Ryan of Social Elf @social_elf - VA of the Year for North West England 2018 presented by #NWVAConf18 host Joanne Hawkins @executive_vpa & Andrew Jardine @IAM_1915 https://t.co/Sx72xEXCW1	0.566	0.048
· Solar news June 22, 2018 at 04:41PM : https://t.co/PmEsoP9CuV #pv #reneweable #solar	0.566	0.048
· Casioís Eccentric Product Culture, Built on Embracing Failure (my 2003 article, in appreciation of cofounder Kazuo Kashio, who died this week) https://t.co/aF6VYCEs8f @medium #productdesign #product #leadership	0.565	0.048
· NYís federal primary is four days away! Join us Monday night to phone bank in support of @dana_balter: https://t.co/pzBzTbAKcS https://t.co/ohPbnvsHqR	0.565	0.048

NEGATIVE Dimension 2

· @Willie_Beamerr I didnít get to watch it but hey a win is a win [emoji]	-0.445	0.03
· @markabaka hahahaha used it metaphorically bc i'm not familiar with those names [emoji]	-0.446	0.03
· @TBoneWFNZ But that doesn't make any since. I mean its almost like picking this guy because you must making a pick instead picking someone who could maybe make the team this year.	-0.447	0.03
· @tgodekjr I did do background screaming, but I think that is just me looking stupid ;o)	-0.447	0.03
· idk how long ive been crying but they all look so beautiful	-0.447	0.03
· @kthechosenone yes lol... and it didn't even fit well. he's gonna be upset with himself about that.	-0.449	0.03

· @titoflo1327 @alexiss_maray Hell yea don't think a day went by that I didn't get roasted	-0.449	0.03
· @SwamiGanesan1 Ok, point taken. But if you're going to make this about race when it isn't about race, we have nothing further to discuss. Bye.	-0.45	0.03
· @so_influential Oh no! That is not how we want you to feel. What's going on? I want to help. -Tiff	-0.45	0.03
· @_salOxo I know sooooo soon I cnt believe it	-0.45	0.03
· @tubirfess I'll be like " hey im totally cool with it, but i may ask u many question about it because its too rare happen in my life, but if ure not comfort with it just tell me and i'll stop asking"	-0.451	0.031
· @doneljefe @Lanaluart goddamnit these are so cute	-0.453	0.031
· @jenevieve22 I love to be holding them they are very damn beautiful	-0.454	0.031
· @HiruniDissanaya You know how I feel [emoji]	-0.454	0.031
· @karlamata98 You really right though, i love ya [emoji]	-0.454	0.031
· i dont trust that a couple members are blonde like that i feel like theyre gonna dye them another color :((	-0.455	0.031
· @LitSego @psndaba I really hope you are right. He is my fave [emoji]	-0.455	0.031
· @castycue me i don't know so i av no comments	-0.455	0.031
· @xSamuelSzx I feel bad but he doesnít seem fazed, any normal person would just make a new account I mean you donít wanna be known as a nonce do you	-0.456	0.031
· @RichiTwoshoes @DavidWilletts3 @HandlebarWisdom @DefenceHQ Ok - so what evidence do you have that proves this to be incorrect ?	-0.458	0.031
· @NerdyAndQuirky I'm sorry :(	-0.458	0.031
· @cuhmilee i know thatís why i said it b	-0.46	0.032
· @outrotins ayooo!! ahh its okey,im a mulfand too! [emoji]	-0.461	0.032
· @beccas1434 @wesley_jordan [emoji] [emoji] Iím so happy for them. [emoji] [emoji]	-0.461	0.032
· @moosemousse @realpaullynch @ChristieElanCan Ok have an x.... why are you x asks the border guard erm i would rather not say... ok turn around. Or we could have male or female. Anyway im not arguing with you anymore atleast i know when i start to identify as a squid an wanna marry my dog your will be behind me	-0.463	0.032
· @Blvckmonk She doesnít have a twitter but Iíll let her know	-0.463	0.032
· lmao you just made me realize that i am really ugly	-0.463	0.032

· @Abbeylou14 @RoadTripTV Aw well I hope u enjoy urself and stay safe x	-0.464	0.032
· @heylakateeng You're welcomeee love you too [emoji]	-0.466	0.033
· @beccadavitt God Iím so jealous I donít mean to brag like but Iím in Vietnam but Iíd still love a chines	-0.467	0.033
· @realDonaldTrump I don't really care do u ?	-0.467	0.033
· @MaxAtkVGC nah he's bad	-0.467	0.033
· @NeilJackson10 I'm quite sure...	-0.469	0.033
· @CNN What's your point. Color is irrelevant when it comes to good officers but I realize that doesn't advance your race baiting narrative.	-0.47	0.033
· @bmacdsst @thistallawkgirl I hope youíre right.	-0.47	0.033
· @snowberrytae Um hello I'm here to steal this thank you	-0.471	0.033
· @SectyHarris @activistliberal @4thfloorview @Phil_Lewis_ @HuffPost That's funny, I don't see where anyone said that at all. Oh, and for the record, kidnapping is illegal. No matter who you kidnap, it's still illegal.	-0.473	0.034
· @OfficialBaileyM Yassss [emoji] we love you too!!!	-0.473	0.034
· (also please rt those if you feel so inclined itís important that we donít boycott a female hero movie!!)	-0.474	0.034
· @YuuKoitoxTouka @JiKxxn im not gunna post it here anyways dont even try lmao	-0.474	0.034
· @LMStewarty I was thinking the same thing..Glad to know its not me alone nah [emoji] [emoji]	-0.475	0.034
· @vi_anaa @SiahhLaw @jtmejia_ @RedCarpetRich i wonder what they be having then lol	-0.475	0.034
· @jacknbridge @MAGAlover99 @SandraTXAS @joshdcaplan It would be nice to find out who they really are!	-0.477	0.034
· you donít know what itís like !!!!!!!	-0.477	0.034
· @adam_253 Dude I don't care what you say that's pussy shit. Have fun shooting that if you ever get the chance though.	-0.478	0.034
· @mintaellaa omf i didnt realize	-0.483	0.035
· @don_jidz If this is how man on here are kicking ball I don't want it bro	-0.484	0.035
· @softkz_09 What really :'(fina cannot meet faz there la mcm tu D: I'm sorry for cannot helping :'(	-0.484	0.035
· smh i guess iíll go cheat since you want that so badly [emoji]	-0.484	0.035
· @warriorAndorian Yes I'm sure	-0.486	0.035
· @Jordan_miggy LOL I am so dead! How could you do this?! [emoji]	-0.487	0.036

· @NcFortwiter I'd be pleased if you do. Rephrasing is fine too.	-0.49	0.036
· @KingPeyso @BriThisBriThat [emoji] [emoji] [emoji] [emoji] you were but itís okay. Still love you.	-0.49	0.036
· @ZJOIN85 @D47372901 @PaulMalignaggi I still donít understand what a ìside piece ì is. Lol. If itís a girl on the side why did he care so much? I just figured Iím to old to understand [emoji]	-0.491	0.036
· @spiritedlunakat Iím so happy I got to see him live [emoji]	-0.491	0.036
· @NEPTUNYOON omg goodluck! when i hacked it it sometimes didnt work but i hope it does for you !°	-0.492	0.036
· @KiingEspy And if I think sheís cute we both bout get some [emoji] [emoji] [emoji]	-0.492	0.036
· @gabsomfg @ItssName Donít worry, itíll be like youíre there ;)	-0.495	0.037
· @TyCrime @balvertos @FOXSoccer Nah itís better if he didnít	-0.495	0.037
· @karna_sakthi Hey I seen this, but not aware itís u. Great [emoji] [emoji] [emoji] [emoji] [emoji] [emoji]	-0.496	0.037
· @LiamPaulCanning I can do that its not hard	-0.496	0.037
· @MRNIKSTONE ew i would be so disappointed	-0.497	0.037
· @desicheshire hehe im glad you do! i hope we win	-0.499	0.037
· @_mansa Im so sick [emoji]	-0.499	0.037
· @CallMeDadddy Lmao I'm bs cuz I'm bout deaf when I say huh I really ain't hear u	-0.5	0.038
· I don't like being around you because I feel that you make me feel bad	-0.5	0.038
· @Iliana_dh Lol I donít even know what that is	-0.5	0.038
· @143Karasage Im fucking invincible	-0.5	0.038
· @Sydneykristine Itís sad :/ but now I can take you out ;)	-0.504	0.038
· @snowdrop284 @FLOTUS @SecAzar @HHSGov Hi, I was extremely perplexed by the jacket. Couldn't figure out why she would wear it. Maybe part of me just doesn't want to believe that she doesn't care.	-0.506	0.038
· @erica_shay_key I know, Iím not ready [emoji]	-0.511	0.039
· @UzoForGod Wow weíre fried [emoji]	-0.513	0.04
· @YOONGlSSI Shit damn u rite fam I mean I'm practiced enough but you can use baby pliers they'd prolly be better	-0.514	0.04
· @optimus_hwangmh It's Japanese, but okay as long as you understand [emoji]	-0.515	0.04
· @yinzerdunny @thehill Oh my goodness you are just seriously dumb lol no one is saying that but	-0.516	0.04

you will never understand that so Iím not going to waste my time explaining basic logic		
· @Ashton5SOS I love you so fucking much I think I'm going to cry	-0.518	0.04
· @Niishatk You seem so proud, i wouldnt be [emoji]	-0.519	0.04
· @slayvocals well i do so oopsie it is possible	-0.519	0.04
· @bigd6777 @isaidayeyooooo @MATHHOFFA @Tsu_Surf @hollowdadon @CHARLIECLIPS @MRDIZASTER @ored973 Nigga... who the fuck are you to tell me I donít know battle rap because I donít like what you like? You niggas turn goofy when somebody doesnít have the same opinion as you.	-0.522	0.041
· @GodlyExecution ì youíre right but where we are going Achilles ?!, do you know any places seal food ?.îShe sweats as she chuckled a little bit ì well I ate before, and Iím not hungry I will just buy some snacks.îShe said to him and waiting for him to walk	-0.526	0.042
· @Kim_xD93 @AmazingPhil ITS SO GOOD I DIDNíT KNOW IT EXISTED	-0.526	0.042
· @thea_lim I squealed when I saw it! Hahaha. Itís delightful.	-0.527	0.042
· @astrodreamergir Ah but this is why I love them	-0.528	0.042
· @UrplePing0 "aw heck kid i just want to be friends you have so much life to look forward to so much to learn ha ha i'm jealous"	-0.536	0.043
· @mhairi_97 @maddiedagg_ Ps. Iím extremely jel [emoji]	-0.539	0.044
· @AjnaPapi I hear you. I feel like itís real simple, itís not that complicated but to each is own lmao	-0.547	0.045
· @cherryyunhyeong You don't know it but... as i'm small i'll hide in your bag!!! Jajaja	-0.555	0.046
· @IQ_Adventures @WadjetEyeGames @GrundislavGames I may try it soon. My game has some very unique features for an adventure game so I hope my ideas don't get "copied", that's another concern :p (though they're so tough to implement I doubt anyone else would dare lol)	-0.559	0.047
· @_notbabe noooooooo [emoji] [emoji] itíll be over before you know it ily	-0.568	0.048
· @prompt_SD @hibbsforchange @GowenAnita @CNN Well I actually do have outrage for that. And I don't know you so it would be hard to have outrage. I also have outrage for this.	-0.569	0.049
· @XeniaKaepernick Thank you!! Iím loving it here so far [emoji] [emoji]	-0.572	0.049
· @0H0UR1 Trump being fucking terrible and people recognizing it doesn't mean his crimes should be forgotten. You know thats not how our justice system works, right? It's important that you know that.	-0.575	0.05

· @yernaizu [emoji]! I think I did terribly but I'm glad it's over haha.	-0.579	0.05
· @BeautynDesire Oh bitch let me snappppp youuuu rq lmfao but I jah donít /: fr lol I guess I need to spread out cuz itís clear your here [emoji] [emoji] [emoji] [emoji]	-0.591	0.052
· @wittycheese_ @ArcticMonkeys Yeah it's very me that's why [emoji].	-0.592	0.053
· @Calliethulu no honestly, I'm sorry that must quite suck :/ *hugs if wanted*	-0.596	0.053
· @retroxirwin if you can help me thatíd be amazing if not Iíd understand [emoji] [emoji]	-0.615	0.057
· @TomArnold Lol haha your so stupid! Didnít you know if anything exists MuleíR has it. And guess what that means when evidence is sealed? Youíre going to fall... you believed him [emoji]	-0.616	0.057
· @DougiePoynter @TomFletcher @harryjudd I'm so glad you didn't killed them [emoji]	-0.618	0.057
· @btykiwi ill spam u dont worry ! and its ok letís just support our faves tho thatís why weíre here but yeah i feel u ìsometimesî [emoji]	-0.623	0.058

POSITIVE: Dimension 3	Coord	ctr
· TYLER JUST SHOWED UP AT MY WORK WITH COFFEE AND BREAKFAST FOR EVERYONE AND IF I DONT MARRY THIS MAN PLEASE SHOOT ME HOLY HELL WHAT DID I DO TO DESERVE HIM	0.672	0.12
· @seekingBushra How about you just hop off my ass and stop bashing me for voicing my thoughts? If you donít like my tweets that much, then stop reading them.	0.651	0.113
· @CarlBovis_AFC Some of the reply you get here..smh, they dont understand that you want to pointed out the negative bias towards wenger in the past, for years ìtheyî criticized him for something that now ìtheyî consider ìsmartî and ìbrilliantî move eg.golovin, and they said ìlet it go carl..îbs	0.631	0.106
· Donít allow me to explore you if we are just going to end up like strangers	0.606	0.098
· Dont wait for me to die for you to realize what youíve done/lost	0.606	0.098
· @FearIessCourt @snakelor_swift @Chriscqma @IamSeruzna @ComplexMusic Look. I donít care what yíall think, I stan someone who are using their platform to raise voice for women empowerment and lgbtq in a conservative society like South Koreaís. I can stan whoever the fuck I want to, so stay pressed and salty	0.562	0.084
· @bangstan7 @BTS_twt ARMY [emoji] PLEASE help me report this video please https://t.co/xU9bpG8nMh i didn't want to do the interview now its out	0.561	0.084

· Iím impulsively tweeting my thoughts more bc I want the people who donít truly fw me to unfollow. Leave me alone	0.549	0.08
· @izneerehanna Hahahaha. I just do my own work. He still doesn't believe me. Sigh.	0.548	0.08
· @CulmoJ Don't worry I found out I just hated myself	0.546	0.079
· Don't miss out on your chance to win! Visit us on https://t.co/iBjVAOlxNK and check out #FrontierFunFriday!!	0.545	0.079
· so i found out my dog died like a week ago and the family (cousins) who adpoted him didnt even inform us [emoji] fuck such a bad day why!!!!!	0.533	0.076
· Charlie Vox charlievoxofficial singing Wanna Be Startin Something-LIVE @ SOBís Donít forget to Download Charlieís Summer HIT [emoji] [emoji] [emoji], ìYou Should Let Me Love Me Youî available on ALL DIGITALÖ https://t.co/L9FReJ3LrM	0.531	0.075
· baby hates when we walk by a human and they dont say hi like she just followed this man in the opposite direcfion for three minutes and i had to ask him to stop, turn around, and greet my dog so we can get back home	0.529	0.075
· @Pru_UK Actually yours was resolved next day. Just don't understand why you write to thank for docs but send docs back separately (and don't mention they're coming back). Inefficient for you and confusing for me.	0.517	0.071
· I won't be watching this game because (a) Nigeria is going to win (b) my wife don't like me shouting over a goal.	0.512	0.07
· @birbigs But I donít want the rest of his cast around either. They signed up to star with him.	0.51	0.069
· Okay so the email that I got that Daddy Yankee was coming to Chicago wonít pull up the tickets. Did they lie to me? [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] [emoji]	0.506	0.068
· ìI told em 2018 ainít no holding back/Iím bout to blow up @Royceda59 until he phone me back, bout to call up @iamKingLos like let me hold a track/I talked to @CyhiThePrynce last week, I can show the factî - @Anthony14rapper #PowerMovesOnly	0.505	0.068
· Tell my son come to my [emoji] & leave me a rose [emoji] stand on all 10s and don't never put your trust in hoes [emoji]	0.505	0.068
· Hey, I have opened my own online store. Do check it out and let me know what you think? https://t.co/6EeRhARyPh	0.498	0.066
· Amend that: I donít think black men should be allowed to speak about ANYTHING in public without first studying black womenís work.	0.497	0.066

· [I donít even cry, BECAUSE jungkook started my sex drive and who did I say I want to dick down??? Hmm? YOU https://t.co/d6Z59Qteyt	0.494	0.065
· You sir need to hold your own against these anti-American Democrats. Do not compromise, do not bend and do not give in to their antics. Build the wall, end chain migration,end sanctuary cities - protect us !! https://t.co/g2Oo41gnpT	0.492	0.065
· @Safaricom_Care Then give me 8 points for the cash prizes or the apartment. Don't give me 8 Ksh. worth of airtime. From the transactions that have made me earn those #MaishaNiMpesaTu points; you can see why I think I deserve more.	0.492	0.065
· We are 1 week away from our 2018 Summer I.D. Camp! Donít forget to sign up by clicking this link: https://t.co/1YZSZHIHh1 We have an awesome and talented group so far! Look forward to seeing everyone soon! [emoji] [emoji]	0.491	0.064
· Hey other black folx:I'm not asking you to fight for other marginalized groups if you don't have the energy or if you feel some kinda way cuz other groups don't fight for us. Fair.Just don't use white supremacy tactics to justify why you ain't fighting. https://t.co/TmaiX0jvYO	0.49	0.064
· @SkywayOctane Mine has always been weedle because all the boys gave me those cards when I didn't know how to play in the third grade [emoji]	0.49	0.064
· Tell me why my eyes started to water seeing jungkook with red hair? AND HE AIN'T EVEN MY BIAS!!	0.488	0.063
· @PMOIndia @SushmaSwaraj @HMOIndia I don't know who else I tag but I just stunned to see this vedio those are possibly from haridwar as per social media ...Sir/ ma'am please please have look at this vedio till and try to feel the scream of a girl in between..Please help them sir https://t.co/6g3lnQHfom	0.487	0.063
· I wanted to get on here and drop some knowledge on yíall, but the end of his tweet sums up my point!!! A man paying the bills doesnít mean the household is 100% dependent on him. https://t.co/9WtLUuuxsH	0.486	0.063
· I feel like @terrycrews probably sensed this disturbance and cried out, but didn't know why. https://t.co/ZwuqbHPN6m	0.484	0.062
· @ABC @CarlosJosu1 Look at them look at him like ìyeah right, like we want more brown peopleî. His administration didnít even help Puerto Rico after the hurricane, they are still suffering.	0.483	0.062
· I donít understand why people love to stay in the fast lane on long stretches of highway gtfo my way if youíre only going 85 if you see me coming in at 110 [emoji]	0.483	0.062
· The most Philly ass shit happened to us yesterday We were walking down south 2 dudes skateboarded by	0.482	0.062

and asked Kyle where they can get a good cheesesteak (we had JUST left our fave cheesesteak spot on South) and seconds later someone drove by blasting meek mill [emoji]Luv u philly		
· @DJDaymos Dont worry you just ignoring what i am saying now lol	0.481	0.062
· I wish she did hang up that phone on me !! I would of Uber my ass to that leasing office so fast!! COME FIGHT ME BITCH YOU WANNA HANG UP!! COME OUTSIDE!!!	0.477	0.061
· Someoneís Dominican titi staring me down like I donít like mofongo, she better relax	0.472	0.059
· Are promoted ads on Twitter targeted? If so, the algorithm doesn't seem to be working. No 'About Us' page but here's what I found in their past publications (which only totals under 2 dozen) https://t.co/nxobgDs96d	0.471	0.059
· @brijeshkalappa You want me to take Saifuddin soaj name n wat he said ???	0.469	0.059
· Hey man, do want all my treasures? Go to find it if you want, Iíve sent them all on the map.My referral code:15GaPP https://t.co/LvTSH7LJEZ https://t.co/H6jqSF2cnG	0.468	0.058
· I really luv fashion my girl asked me to put her an Dinner outfit together...What y'all think ? https://t.co/ouyNce3Yvo	0.465	0.058
· Who got more records than me right now?!?!?!?! I got 55 record for one tape i jus did those in 2 months....dont let ly shoppin habit fool u ill outwork all u niggas	0.462	0.057
· Cover Reveal $25 Giveaway - Check out the cover of The Songs of You and Me by Mylissa Demeyere and enter to win! https://t.co/L1BhS9yRWS	0.461	0.057
· I acknowledge you agree with me on what I presented & I appreciate it, honestly I do. Just pointing out you keep making false assumptions of me. Trying to show you my character is all. If you want elaboration on something then just ask & don't assume about me. https://t.co/HNMUxzE7M5	0.461	0.057
· @jezebeIla @savuhna @HannahNiicole13 ...donít even need to say his name do I	0.461	0.057
· @ant_shantt @Namrataye Even if he tried talking to u 18 yrs back i dont think u were as intereted in him as u were in barbie [emoji] [emoji]	0.461	0.057
· If you pay for everything in a house where you are living with another adult you gotta start telling them to not keep on any lights cuz they running up yo bill	0.456	0.055
· @tamiaaaa__ ìLet these hoes knowî. Girl the hoes ainít gonna want him after they see the d trash. She didnít make one sound [emoji]	0.456	0.055
· Can Somone explain to me why Alex Iwobi is not on the field pls, I donít understand?	0.455	0.055

· @ShrinkGov John that image didn't move me. I know what is happening. I listen to the testimony of the lawyers who are trying to secure rights for the snatched. I don't play the monkey poo fight at the zoo that the propaganda makes left and right white engage in.	0.452	0.055
· @miscusername2 @donwinslow @realDonaldTrump The first part of my tweet asked for his sources, because I want to believe him. I never said I believed the daily mail. The second part of my tweet says I agree with him, people should not believe everything they read.	0.452	0.054
· please donít leave me hanging here alone	0.451	0.054
· Please dont count me in mourners for Charles Krauthammer. Hammered Kraut was a deceitful lying hate filled War mongering neocon. When i see his face i wanted to punch it almost as much as Hannitty. He was leader in Black/Muslim lies against Obama. Where Trump got his birther BS.	0.448	0.054
· @thehill @stevemacwv @SecNielsen lied to everyone on national TVShe deserved thisShe&Trump&Session now want to lie their way outDON'T let them	0.446	0.053
· Girls who donít keep their toe nails painted scare me	0.445	0.053
· @lyndamk I usually do bring home baked cookies to Council. Watch for them at Council II!	0.444	0.052
· @realDonaldTrump Just because you lack the ability to feel empathy, doesnít mean the rest of us do.Nothing phony about our outrage. We will stop you.	0.444	0.053
· i lied i didnt go to sleep apink has awoken me	0.444	0.052
· 17 Small Mistakes That Stop You From Getting Big Time Traffic To Your Blog - Don't Make Any Of Them https://t.co/dhVqEdObyN https://t.co/Qf5GrR9PNQ	0.442	0.052
· @Lizerenity @MichaelAvenatti @AmericanAir Send everyone of them back to their own country let the leaders there deal with them no more entry at our border not one ILLEGAL ADMITTED	0.442	0.052
· @Nadallica86 1)How did she persuade him to do this,I will never know.2)Uncomfortable.3)We will lose,don't worry.We know shit about football.I was joking. https://t.co/iJtursXPP4	0.442	0.052
· Check out my book - 'Don't Let Her Wake Up Alone' - on #BookBuzzr - https://t.co/SxtYjVVWdm	0.441	0.052
· I keep getting weirdos request me on Facebook [emoji] I donít approve anyone I donít know on fb	0.441	0.052
· i donít trip over what people talk about on here cause yíall talk about stuff yíall know NOTHING about	0.439	0.051
· @unlovablefuck Definitely them. Don't let them lead your life	0.438	0.051
· @angelchrys Yeah itís just frustrating how fanfic gets dragged by people who genuinely have no understanding	0.438	0.051

of it at all. I havenít seen orbs in so long. Almost makes me wanna try to write something in one of your fandoms just to see if you find it. Who you reading these days?		
· @shayjanae13 @Shay7Eleven @wockstvr @fletcherjohniqr @skipgive5 So all we have to do is stop talking about it and it will go away? Shit someone should have told me that years ago.	0.438	0.051
· donít touch me	0.437	0.051
· Don't touch me	0.437	0.051
· don't touch me	0.437	0.051
· Dont touch me [emoji] [emoji] [emoji] [emoji]	0.437	0.051
· @davidjacobsonbk @ryanschott @Patsme @FloydFlanders @OneWheelProds @DavidWSTX @VeeVee @Tharp1965 @TomiLahren He wanted me to show you this since you blocked him. https://t.co/OM6Gynnowj	0.437	0.051
· Don't forget who helped you out while everyone else was making excuse .	0.436	0.051
· Sooner or later the iHub basher morons are going to realize putting up an ask they don't really want to sell in an attempt to fill on more on the bid doesn't work when whales are involved who don't post everything on irrelevant forums on iHub. $DIGX $IJJP $HIHI $TMPS $INCC $EVSV	0.435	0.05
· Everytime someone wants to travel they want me to come and do all the planning well i donít mind I love free plane tickets [emoji]	0.435	0.05
· @mainameiskai SZN is coming sooner than y'all think. Don't say that I didn't warn ya!	0.435	0.05
· No new video this week! Life happened and time got away from me, but I have lots of great things coming!But you can always check out my other videos! https://t.co/mdWluSAnk0#amwriting #authortube https://t.co/1J6lx1Em8R	0.434	0.05
· I get dmís from men regardless of status. If I donít want to talk I simple donít speak. I also filter my dmís.	0.433	0.05
· Don't kidding me!! #Azusa	0.431	0.05
· @bangstan7 @syjeonn @BTS_twt @BTS_twt we are begging you to kill us and give us a heart attack please. 14 min left hopefully they will post something	0.428	0.049
· @Just_Vodka8 You had me at 'don't enjoy watching baseball' and lost me with all the babbling after	0.428	0.049
· @PatrickRandall @TrueFactsStated Speak for us - and I will hope fir your safety - we need your voice	0.427	0.049
· @kriktsune are you telling me he wasnt already https://t.co/1iCOlRWxNR	0.427	0.049
· @SpockResists True some of them are "chaos agents" but not all.Indeed, if you don't know the person already, and if you don't carefully vet your new followers	0.426	0.048

you shouldn't join in.They can serve a useful purpose in joining like minded people to #Resist.		
· lol this dude with a girlfriend keeps sliding into my instagram DMs trying to flirt with me and essentially just called me a ìside slam-piece.î you donít even deserve a girlfriend, boy bye.	0.425	0.048
· donít touch me to pisando f O fO https://t.co/Ak9juu7JKX	0.425	0.048
· @everhopeful1000 @nadhimzahawi Do not know how to make the ordeal heard! I tried to ask for help but was met with silence. May be we need to put up an act like inside Dorchester to get heard!	0.425	0.048
· @achrisvet @LOLGOP Well they dont want to be sent back They claim "Assylum" which is bs	0.425	0.048
· You don't know what it does to meReading trade rumors constanltyI'm haunted by Blues historySo give me a trade done magnificentlyTrades, trades, I want tradesThere's no substitute for tradesTrades, trades, make some tradesThere's no substitute for trades#stlblues	0.424	0.048
· likes & retweets don't do it by your self just tell me	0.424	0.048
· @TlFPHANY do u want to add me!!	0.424	0.048
· @imamrapalidubey Well said @imamrapalidubey ji kindly ignore him and just think one thing we love you and keep doing such amazing movie ..#bharatmatakijai	0.423	0.048
· i DON'T EVEN REMEMBER WHAT CALM DOWN MEANS AFTER WATCHING @BTS_twt on #BTSxLotteFamilyConcert	0.423	0.048
· @bts_bffs DON'T MOMMY ME	0.422	0.047
· I donít need anyone, they all need me.	0.419	0.047
· I don't know what the fuck I did to my voice, but whenever I try to sing I get an itch in my throat and cough	0.418	0.047
· We will win the game today by Godís grace! [emoji] [emoji] @NGSuperEagles #WorldCupRussia2018 donít disappoint us	0.417	0.046
· @Scotttaylorva let me get this straight. You want us armed with military weapons to prevent tyranny. But you deplore peaceful protest against tyranny?Guess it depends on whose side youíve picked. You picked evil, sir. You will not be remembered well. https://t.co/eSIYuO9eXU	0.417	0.046
· @GeorgeTakei Wow and I didnt think I could be more disturbed by all this	0.417	0.046

NEGATIVE: Dimension 3

· @steve_vladeck Also, FWIW, it's hilarious that Dalmazzi is no longer one of "THE CONSOLIDATED CASES."	-0.657	0.115

· @WonderBread941 That's a flaw in the system, like saying good service is perfection. 4 out of 5 stars should mean really good but keep striving to be better.	-0.609	0.099
· @MagesticSalah @The_Koxy @Thfc_Scops @JamesPearceEcho Coutinho is so fucking good. Heís a top 10 player in the world, dude is lights out.	-0.548	0.08
· @vineet_red @ThomoJosh @SimplyUtd But Fred is one of your newest signings while Sturridge is most likely on his way out...if that's his quality then that's just underwhelming lol	-0.545	0.079
· @bobpockrass Possible that it's not the driver but the car? Kenseth is so much better than his finishes this year.	-0.54	0.078
· Being suicidal is liberating. Like the worst thing that can happen is death, and you're okay with that.	-0.534	0.076
· #MiraclesOfGodKabirHindus say that Ram is great, Muslims say that Allah is great, Christians say that Jesus is great & Sikhs say that Nanak Ji is great.†Truth is that Supreme God is the greatest who's the father of all. https://t.co/nyYdZvDhpd	-0.526	0.074
· @MhelFoxx It looks soooo good. Good morning Mhel, and have a great day.	-0.524	0.073
· @N8Sutcliffe It must be past the sore stage now. Itíll be so big it creates an echo.	-0.52	0.072
· @democrat_donald @FoxNews @realDonaldTrump Keeping tax returns private is fine. This isn't. Why don't you do some good and retweet this instead of whinging abouyv tax returns 18 months later. It's moronic. [emoji]#IGReport page 294 CRIMES AGAINST CHILDREN Clinton FoundationRape?Trafficking?Worse?	-0.518	0.072
· @MartenJimi @matthewasears @RonaldDPotts1 @NewWorldHominin As an internet lawyer, I'm sure you're aware that each province has its own law. Is that from Ontario?	-0.513	0.07
· @petouyou @CryptoDonAlt German engineering at its finest	-0.507	0.069
· because its really impossible working during the tour, august is the max predicting for the 2nd comeback	-0.506	0.068
· 111 Ave reduced to 4 lanes for construction. Rush hour traffic still moving fast. It's a sign this road is too wide. @bevesslinger @yegplanning consider converting outer lanes to bus lanes, parking, or wider sidewalks. #yegcc #yegtransit #yegbike https://t.co/nXoHeKajC2	-0.502	0.067
· Attraction marketing is the most effective way to bring in prospects and customers to your business, it's so productive that it's the only real strategy that you need to profit IF you stick to these principles.. https://t.co/fvv85I35rw #NetworkMarketingSponsoring #marketing	-0.498	0.066
· Time slams family separation policy with new cover....*This might make a much better story/cover...IF IT	-0.498	0.066

WERE TRUE [emoji] [emoji] [emoji] https://t.co/fEXe83EhxV		
· @staronline Its ok. It is just the "Joke of the Day" by The Star.	-0.497	0.066
· If thatís true, itís even more embarrassing - watching a marital drama play out with the highest power couple of the world as if itís just a common reality show. #growup https://t.co/UAA7KjN34s	-0.495	0.065
· @ChefCruick Never enough is such a powerful song and its the best in the film you should watch the live performance of the original singer it's amazing	-0.49	0.064
· @ChrisGBurns @ECFCJJ @TheRedRoar @BenPBradshaw @Nigel_Farage @campbellclaret Airbus is the major employer offering skilled employment in North Wales. This is why its job vacancies and apprenticeships are massively oversubscribed.	-0.488	0.063
· Way better than wine. "a taste of honey, tasting much sweeter than wine" https://t.co/BspboseVqy	-0.487	0.063
· i've darkened this dp. the original one is way much brighter.	-0.486	0.063
· @ShowcaseUS No thanks. A picture is worth a thousand words. And I've already paid you, so maybe you should do any further investigating work yourselves.	-0.486	0.063
· It's Fandangle time! @MariCockerell talked to this year's director on how the show has changed and is much shorter than years past... https://t.co/pWTppGtwlg	-0.483	0.062
· Erm @Eptic youíre crazy dude [emoji] this is insane https://t.co/TVepE5468V	-0.483	0.062
· 2019 Chevrolet Blazer, Fatal Uber crash update, Electric-car tax credits: What's New @ The Car Connection - https://t.co/KzE1fzmxVC Chevr... https://t.co/RyP8yzVtD1	-0.482	0.062
· One last thing, it's too victorious for a midpoint. Star Wars was modeled of The Heroes Journey, and the midpoint of that journey is supposed to be our heroes' darkest hour. The plot beats in TLJ resemble this (The Resistance being on the run and eventually reduced to nothing)	-0.482	0.062
· @leedqin For now there is just a long angry paragraph about summer, but thank you, youíre lovely [emoji]	-0.482	0.062
· @_DollMaker @bleudawn7 @jolievie123 @soonthareeya @Branka_Malle @Lalla584 @witch66fr @kimsippingtea @LamissMchat @sami55832 The anemia and all I already had it cost well is chronic and I've been fighting depression for almost 2 years but this wrong medication made everything worse	-0.48	0.061
· @FatOrangeIdiot @joseiswriting @real_farmacist Youíre right. There is no justification for fake news.	-0.477	0.061

· Nigeriaís kit is better than the actual players! #WorldCup #NGA	-0.476	0.06
· @glads1951 Hey Glynis, really that just means there is less to share! Kyrie	-0.476	0.06
· @Nanuya Yes it was so funny [emoji]Iím so sad itís the last series	-0.476	0.06
· Kiawah, youíre awesome. But, solid deck of altostratocu, youíre the real MVP. https://t.co/6JR1TXEPK0	-0.475	0.06
· @jinKissLetsgo Hachi this is so beautiful!!! [emoji] [emoji] [emoji] [emoji]	-0.475	0.06
· Iíve already conjectured itís because YouTubeís margins are so low, theyíre forced to cater to advertisers & has no resources left over for moderation. Thatís the optimistic interpretation	-0.474	0.06
· @ArielGZB but their stylings on her were always the best. it's okay she be on their covers someday	-0.474	0.06
· @mitchellreports @Christi22657596 He is by far one of the most uneducated people in this country. He's an embarrassment to society.	-0.472	0.059
· This is very good. https://t.co/Adm3vfntNI	-0.472	0.059
· Man this is so sad [emoji] Rip Mr.Harris [emoji] https://t.co/4PferVCg3P	-0.471	0.059
· Not sure what @henrywinter has seen in Sterling playing for #ENG? He looks good when playing with Suarez, De Bruyne and Silva. Not so much when he is the number 10. It's not like this is his first #WorldCup This is his third tournament. Two goals in 39 games is not good.	-0.47	0.059
· Itís great seeing more organizations getting into competitive splatoon. Good on @GhostGaming_GG. This game has a huge scene thatís only growing and i really expect big thing from Nintendoís brand into Esports. Mario Kart, Splatoon, Smash, and Mario Aces are all competitive.	-0.47	0.059
· @HOLYHANEULIE pats your back. it's okay, haneul. you were so brave back there.	-0.47	0.059
· @Tvrrell @wowjymon its so delishhh	-0.469	0.059
· Lordship Road, London. £2,100 PCM//£500 PWï Recent Refurbishment ï 2 Large double Bedrooms ï 2 Bathrooms one of which is on-suite. ï Virgin Cable TV. All Sports, movie channels. Super Fast Broadband internetï Furnished to a High Specification inclusive 42 inch Smart TV https://t.co/H3DpFjFLEH	-0.465	0.058
· this entire thread is centrist liberal nonsense of the highest degree https://t.co/0P6WMNg1KM	-0.463	0.057
· @Maryam_Rajavi ya, that's ok.	-0.461	0.057
· Omfg Narwhals, unicorns of the sea, you [emoji] magical [emoji] creatures you. The underwater world is all the more special with you in it.Iím sorry my idiotic species is opening your home to arctic drilling, which will	-0.46	0.057

surely affect your ability to echolocate.https://t.co/kKb88Djh3o		
· Jackson is so dramatic and Iím over it	-0.46	0.057
· #UHIVE Team Success Stories - G Cloud app is protecting over 5 million users around the world, it is a free Android and iOS backup app that is simple and safe to protect all users data on the cloud.https://t.co/THVzyGrt6x	-0.459	0.056
· Apparently God has a quota of two when it comes to great white rappers. The Beastie Boys and Eminem.THEN YOU SHALL GET NO MORE!Ok, the Run The Jewels guy is pretty good too.	-0.459	0.056
· When sheís all happy about those freshly picked cherries [emoji] [emoji] cariniro Internet connection is pretty bad here so itís more like a little digital detox - Iíll be back in the office on SundayÖ https://t.co/IpUsriGqRQ	-0.458	0.056
· Chopsticks. With black sesame ice-cream. @SerendipityAus are you for real? This is so offensive [emoji] [emoji] [emoji] https://t.co/vSUx6rKvbb	-0.457	0.056
· [emoji] [emoji] [emoji] [emoji] [emoji] this is so funny https://t.co/tQ8CHsLcZV	-0.457	0.056
· Vaping360: RT opinion_joe: Which is worse: a teen who vapes or a teen who smokes? Which is better: an adult who smokes cigarettes, or an adult who uses e-cigs? Why is so hard to comprehend? https://t.co/B4J2qhKv9Y	-0.456	0.056
· @tajnalee You look very nice thatís all	-0.456	0.055
· @sdean23 @CarmichaelDave its not the fact that they passed on Doncic, really. It's their justification of taking Bagley over Doncic. Bagley isn't a 3. Taking the ball out of Fox's hands is a bad reason. They like MB over Doncic? Fine, time will tell on that, but their reasoning is bad.	-0.453	0.055
· It is vital that our younger generations, the guardians of our future, develop strong awareness concerning the futility of war.	-0.452	0.054
· Clickhole is undefeated. I've seen so many people think this is real https://t.co/AghvfLZvP3	-0.452	0.054
· this is so fucking unreal https://t.co/gnC5AJrBuK	-0.449	0.054
· It was great to chat with the @mrianleslie and @RSAMatthew about all this, which ranges from the role 4chan played during the 2016 US presidential election to how algorithmic docenting, coupled with personal choice, contributes to ideological entrenchment https://t.co/DCUc7M1aMA	-0.448	0.053
· @PressSec 2300+ Sarah! THAT'S SHAMEFUL!!	-0.448	0.053
· RT NCSC: "Physical security used to be the security we were most interested in, but of course that's no defence against a cyber attack" - Joanna Place, Chief Operating Officer, BankofEngland https://t.co/9dboJZX5jx	-0.447	0.053

· Yo...this is massive...congratulations https://t.co/uyuJb5s2uw	-0.446	0.053
· @iamintrovertt Hahahha..! You are so innocent mind girl. Adv high Court is short for ADVOCATE High Court.	-0.446	0.053
· At FDM, the #diversity of our teams is what makes our organisation so strong. In the words of Maya Angelou, ëin diversity there is beauty and there is strengthí. Weíre so proud of our diverse workforce! #Pride2018 https://t.co/fXjEtw43oJ	-0.445	0.053
· Talk is cheap. Time to slap significant tariffs on Mexico.It's time they get a reality check!#BuildTheWall [emoji] [emoji] [emoji] [emoji] [emoji]#EndChainMigration [emoji] [emoji] [emoji] [emoji] [emoji]#NoAmnesty [emoji] [emoji] [emoji] [emoji] [emoji]#BorderSecurity [emoji] [emoji] [emoji] [emoji] [emoji]#MAGA_45 [emoji] [emoji] [emoji] [emoji]#AllStarLadies [emoji] [emoji] [emoji]#MaxWarriors [emoji] [emoji] [emoji]#RedPilled [emoji] https://t.co/RQcrIGlHNQ	-0.444	0.053
· @outrotins ayooo!! ahh its okey,im a mulfand too! [emoji]	-0.444	0.053
· America needs universal healthcare. Healthcare is a human right. Literally every other free country in the world has it, and are much better off with it. Why canít the states?	-0.443	0.052
· I know itís a lot, but itís so worth the read https://t.co/LfaTh8uAgy	-0.443	0.052
· Dejan Lovren is too dumb to realize that! https://t.co/FByo1ImOX9	-0.442	0.052
· ÌIn my work with the defendants (at the Nuremberg Trails 1945-1949) I was searching for the nature of evil and I now think I have come close to defining it. A lack of empathy. Itís the one characteristic that... https://t.co/DSF5DxI2W5	-0.44	0.052
· They sound so good live!! It's amazing!! [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] @BTS_twt https://t.co/6V8ovkDu0Q	-0.44	0.052
· The instrumental part is so powerful. Iím- https://t.co/B6OH6PYpHA	-0.439	0.051
· @JamesBondRadio I'm sorry guys, but surely the appeal of #Licence is how different it is. Criticising it for being different and less like Bond, I don't think that's a fantastic argument, when normally you have valid points.	-0.439	0.051
· @AliceDreger HAHAHAHA! That's awesome!	-0.439	0.051
· @psankar @Prabhukbala @msathia @krangana One of the reason for rent increase is bachelors working in IT https://t.co/GJUigTPcfQ is easy to share 5*3000 for flat. Same is difficult for families.	-0.438	0.051

· Gents are killing it at the gym, upper body looks good. Upper legs are good. Lower legs look like a kitten heel [emoji]	-0.438	0.051
· @kylegriffin1 And that makes it humane, acceptable, decent? WTF is wrong with these people?	-0.438	0.051
· I love The Office... A LOT... But, Parks and Rec is funnier.	-0.434	0.05
· @Tozendai @Queer_Tankie lol in context is even worse.	-0.434	0.05
· [emoji] "One idea of handling repeat offenders is extremely harsh, but ..." ~ Ned Wicker https://t.co/KRgV4H46oC https://t.co/1AlhVWnS5B	-0.433	0.05
· @alexpberry @BubbaSr556 @DPersistence @realDonaldTrump Your constant degradation is inappropriate! You, not unlike the PALMER REPORT, are a nonissue! You are irrelevant. Your opinion is irrelevant!	-0.432	0.05
· It was too good to be true.	-0.432	0.05
· Hereís fun! 4 hours into our trip to Alton Water for the #greateastswim & weíre only just passing Crewe. The scenic route is now as bigger a carpark as the M6 weíve been to 2 McDonalds and our youngest has been sick. STILL STRONG [emoji] #ididitforthechristie #teamchristie #greatswim https://t.co/Wyi348jQDP	-0.431	0.049
· @DUALIPA The first picture is so funny! [emoji]	-0.431	0.049
· The temperatures were better for the US Bank employees painting at Plum Ridge Park but the rain was not far off. The ladies worked hard to get the shelter and surrounding amenities painted before the afternoon... https://t.co/6rKKs4WxB1	-0.43	0.049
· @friesb4guys92 And an hour normal folks' - like me- bodies begin to produce less testosterone and more†cortisol, which eats up tissue and increases potential for body fat storage. Also depending on intensity, recovery will be harder. But some athletes are capable of doing it repeatedly	-0.43	0.049
· @ThisWeekABC It's anti-family with health problems and new criminals in future!	-0.43	0.049
· @HoarseWisperer Wow! That's good!!	-0.429	0.049
· @Formallyigb @chzbizman No way that's possible [emoji]	-0.428	0.049
· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/I1pEKHa03C	-0.426	0.048
· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/jp9LWy7SQb	-0.426	0.048
· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/QTNSv7SfGn	-0.426	0.048
· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/z2deuEpUxV	-0.426	0.048

· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/K0e4Aedie1	-0.426	0.048
· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/MjBgGzwYgY	-0.426	0.048
· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/DSLGT4z19L	-0.426	0.048
· Your uncharacteristically sharp tongue is less offensive than ... More for Pisces https://t.co/EmGy3eJfLb	-0.426	0.048
· @ChristnNitemare @SarahKSilverman @xeni Wow, dude this is the most unchristian thing ever! You are full of hate, Christ loved all you hate most! WWJD!	-0.426	0.048

POSITIVE: Dimension 4	Coord	ctr
· TGIF #itscalledfashionsweatylookitup and Iíll be in all day tomorrow too - come see us if you want this di*khead lewk [emoji] @ Claw & Co. https://t.co/DJ82bMx2Ht	0.673	0.123
· Breakfast/lunch time! Iím feeling really sick & this is all I could stomach so thank goodness for these crunchy flakes! Now send me to @LoveIsland please! #kelloggscerealdater #hungry #crispyflakes https://t.co/zW9L6kyDKX	0.667	0.121
· IM LIVE! SO PEEPS COME SAY HELLUR! [emoji] [emoji] [emoji]https://t.co/uKt2TTMHxUAlso check out these peeps! - @AyrockMusic @Alyssa_Rxse @chasifrass4 @PROJ3CTGD @_OnlyVic_ @MrSynnn @Joshh9761 @xLeaahh @OrbzTV @KingTylerish @lolgabiz @Elthirlwell	0.655	0.116
· Ugghhhh!!! I'm literally screaming!!!! Jungkook really looks so good and hot! [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] [emoji]#JUNGKOOK#RedHairJungkook @BTS_twt	0.643	0.112
· Iíve been wishing for a comeback of Kim Seokjin in blonde since last year.. and now here it is [emoji] my heart is so happy, im so happy [emoji] [emoji] @BTS_twt thank youso much for this [emoji] [emoji] [emoji] #BTSxLotteFamilyConcert https://t.co/w7OnQBSAqv	0.616	0.103
· Weíre at Road America revving up for #GlobalMX5Cup. Watch us LIVE at 1:20pm CT today! https://t.co/3mgLav12at https://t.co/x4WjWqwS26 #HuntingtonMazda	0.589	0.094
· God, I HATE him. I really really HATE him. I HATE him so much. I HATE him. I HATE him. I HATE him. I swear I HATE him!!!Oh no... I LOVE him [emoji] [emoji]	0.573	0.089

[emoji]@BTS_twt #BTSxLotteFamilyConcert https://t.co/RXSI1DHdxQ		
· @sensation1204 @BTS_twt PLEASE UPLOAD THE HD PICS SOON, IM FREAKING OUT OVER THESE PICS!!!BLONDE JIN IS SUPERIOR!!! [emoji] [emoji] [emoji]@BTS_twt #KimSeokjin #JIN # [emoji] # [emoji] #worldwidehandsome #WorldwideCutieGuy	0.556	0.084
· thank you for staying with me even though iím so hard to deal with [emoji] thank you for being so patient with me when iím being ìsuper arteî [emoji]... and refuse to admit iím wrong [emoji] sorry [emoji] pagsubok lang ëto!! i love you so much, love @JerickAntonio	0.546	0.081
· RT @TVtaboo: I am now available for #DirectChat via #AdultWork.com. Come give me a call! https://t.co/Z6VK2IvzKW	0.541	0.079
· I absolutely loved DEAR RACHEL MADDOW by @adriennekisner! Check out my thoughts in my #review: https://t.co/yf3j46Ieig #BandL #BandLArchive https://t.co/MVBirckYA3	0.535	0.078
· Wow! Just discovered @Mailbird I love it! It's the Best FREE Email Client for Windows :D #email https://t.co/IrOE40tbta via @mailbird	0.53	0.076
· OMG! I love this game live so much!#Cube TV#https://t.co/Mg8ji0ZCHb https://t.co/qVrnvOZpcI	0.523	0.074
· I'm selling Casio Baby G Watch Jam Tangan Perempuan for RM47.80. Get it on Shopee now! https://t.co/MwVUjKeWhj #ShopeeMY https://t.co/AZi8hVnN8Q	0.516	0.072
· IM WATCHING YOU! [emoji] https://t.co/OBvxDEpKce	0.515	0.072
· Opening ceremonies with my fancy team!!! So many new workouts coming out soon!!! Ekkkkkkk! Iím SO PUMPED! [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] [emoji]	0.507	0.07
· Yo, we're checking this out tonight on @WatchMixer! https://t.co/WLpOMlsOML https://t.co/0y32M8wpJh	0.505	0.069
· Kaya im so proud of you Donny! [emoji] [emoji] https://t.co/uzyadKu4pS	0.504	0.069
· IM FUCKING DEAD [emoji] [emoji] [emoji] [emoji] https://t.co/etxsCx5PTe	0.504	0.069
· Esos escalofrÌos que te dan cuando ves y escuchas arte. Pues eso es lo que me pasa con Kim Taehyung. Es puro ARTE! Those chills that you get when you see and listen to art. Well that's what happens to me with Kim Taehyung. He's pure ART! [emoji] [emoji] @BTS_twt #BTSxLotteFamilyConcert https://t.co/8TqcX1UulT	0.491	0.065
· #FabFridayPost @ethannevelyn @kipperscurtains I'm still wearing orange https://t.co/4Wre85OJnl #gunsafety #WearOrange #MomsDemand #VoteThemOut xoxoxo	0.49	0.065

· Jhope this is FAMILY PG concertI feel baptized and im laying in my bed in Europe#BTSxLotteFamilyConcert #JHOPE @BTS_twt https://t.co/KrtyGYUdn6	0.488	0.064
· Bagged my distinctions! I'm so happy !! [emoji] https://t.co/fqBf1VE8hN	0.484	0.063
· ìYeah Iím fucking sureî [emoji] https://t.co/7SzfbeKsOq	0.484	0.064
· I been on...tell me who's gonna take me off....#ALAAC18 https://t.co/ERTmEV5G6n	0.483	0.063
· Love @alphaoutpost every time I get a delivery from them itís like Christmas. #iamredneckdon #makeshithappen #honorablepatriot #americanbeardedman #patriot #america #americanpride #usaÖ https://t.co/lhpmesglvz	0.482	0.063
· I WANT MOREEEE! THATS TOO SHORTTT!!!!https://t.co/KwzOGfPBRt #NUEST_W #JR #WHO_YOU #Dejavu#20180625_6PM	0.482	0.063
· I've just baked 39,347,370 cookies in #CookieClickers2. Grandma will be proud of me!https://t.co/UcTZoFn0gZ	0.481	0.063
· [emoji] Oodles of love for these buttercream peonies by that oh so clever @preppykitchen He just nails it every time! Love me a buttercream cake... Happy Fri-yay all! [emoji].#buttercream #buttercreamflowers... https://t.co/pUlh9xKvpV	0.477	0.062
· #pawcircle surrounds you RT @badgerdastaffy: Embarrassing #FrogLegFriday on mi way dogtor c ow Iím heelin [emoji] https://t.co/Pm3z18AeZw	0.476	0.061
· They sound so good live!! It's amazing!! [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] @BTS_twt https://t.co/6V8ovkDu0Q	0.476	0.061
· We are so excited for the @FRSAExpo next week! Stop by our booth #1223 to learn about our full line of Silicone Roof Coating products. #FRSAEXPO18 https://t.co/5gxSiGI2IG	0.475	0.061
· Omg Iím so waiting for this [emoji] [emoji] https://t.co/pY0IXoKdz8	0.474	0.061
· So excited for today!!! Be sure to catch my live around 1ish so you can see whatís happening!! [emoji] [emoji] Also join me in 15 minutes for a GRWM live. [emoji]	0.474	0.061
· Oh this @utahjazz team how I love them! [emoji] [emoji] [emoji] https://t.co/9hMNlrSBYT	0.471	0.06
· Happy birthday to this wonderful woman that I love so much!! [emoji] @dinahjane97 https://t.co/lI5Lvdte09	0.469	0.059
· @Hyperec_HRS #giveaway this would be perfect as itís my husbandís birthday today and we love John Lewis xx	0.468	0.059
· Hello [emoji] [emoji] to our newest #CreativeWorks partner trinitydccas! Our very own la1cameron got to meet the admissions team and take a tour! [emoji] Weíre so	0.467	0.059

excited to see how this partnership can helpÖ https://t.co/37rqq7snBq		
· Iím really glad that Iím alive to see this https://t.co/dACQvtihEW	0.466	0.059
· yie! love this girl so much. galing mo!!! [emoji] miss you and i hope to see you when you come back here [emoji] @juliennekryzel https://t.co/24lYv11T75	0.464	0.058
· Good luck @MartinMurrayBox for your fight tomorrow night. We're all behind you and will be cheering you on from St Helens [emoji] [emoji] https://t.co/nLnf3Y9QhH	0.462	0.058
· I liked a @YouTube video https://t.co/yzdRcJ8vuI THE JADE RABBIT IS SO CLEAN!! (Destiny 2)	0.462	0.058
· I'm too poor to upgrade [emoji] https://t.co/npNopL7S4c	0.462	0.058
· Send me love or im die! https://t.co/cn21A3MUMS	0.459	0.057
· I know you're ignoring me but I hope you're happy without me. God bless you always. I love you [emoji].	0.459	0.057
· I wanted it to be a surprise but since itís my birthday I like to thank @chillspotlabel for being my manager!!! [emoji] [emoji]	0.456	0.056
· im killing it for my birthday this year [emoji] wait on it !	0.455	0.056
· 173/365 @SHINeeìAnnyeonghaseyo, SHINee-imnida!îI will never get tired of hearing those words. You taught me so much about facing life & how I could be a better person not just for me but also for the people around me. I couldnít thank you enough for everything. I love you. [emoji] https://t.co/2wN3VccTaF	0.455	0.056
· Nothing says #FeelGoodFriday like a handstand on the beach! As the summer hots up, we're celebrating the #SlimmingWorld members who are feeling sunnier since slimming down - like these gorgeous gals, who say they feel amazing since losing 27st 7lbs between them! #FindYourSunshine https://t.co/AWIBCXEcyJ	0.455	0.056
· We're #hiring! Click to apply: Tax Analyst/Manager/Director - https://t.co/Zgpkl6qJHq #Finance #Mississauga, ON #Job #Jobs #CareerArc	0.454	0.056
· It's just over a month until #2018AACC officially open it's doors! Visit the @VitlProducts team at BOOTH 1876 to find out more about our latest products including the Lu-mini & our extended range of Heated Modules: https://t.co/wll28PWK5k @_AACC #lifesciences #labware https://t.co/eIZ78Zuh0b	0.454	0.056
· I'm looking for a buyer for my listing at 5801 Muirfield Lane Chattanooga, TN. Please contact me for details! #HomeBuyer https://t.co/Np9pZuM9DY	0.453	0.056
· I'm so glad we're raping the planet so these guys can get rich. #MAGA https://t.co/7sLVjXgKX8	0.453	0.056

· I'm so proud of you baby @ddlovato #WeLoveYouDemi #StayWithUsDemi	0.452	0.055
· It has been a while since Iíve seen this wonderful error - still makes me want to #backslap #Revit https://t.co/AQoKeExvUv	0.45	0.055
· HAPPY BIRTHDAY @Jessekoolkid321 I love you so much and I hope that you have the best day ever! Thank you for being such an amazing friend and always being there for me! Gonna miss you next year! Love ya! [emoji] [emoji] [emoji] https://t.co/AjpZREon4b	0.45	0.055
· Polos, chinos, watches ... oh my! Itís all here in this weekís #menswear shopping list: https://t.co/REnNTLIqvi	0.45	0.055
· FFFUCK ME IM ACTUALLY CRYING THE H A I R THEY LOOK SO CUTE https://t.co/oW7Z07sXo3	0.449	0.055
· I be feeling so bad bc thatís my twin my Baby George [emoji] [emoji] [emoji](Justice) i cant do ha like that [emoji] [emoji]	0.449	0.055
· LOL Iím still kicking myself for not grabbing Terry and Hicks during preseason for us all to take a pic and send to Raleigh RT @garnerroad: @TrooKing2 You're a successful Garner Road product too - Mr. Executive! LOL.#BuiltByDesign#Bulldog4Life	0.447	0.054
· :,c its not working for me im https://t.co/hLcLKxKAjw	0.446	0.054
· #Repost @teriohagan with get_repost∑∑∑All I want is for my arms to grow, is that too much to ask? [emoji] [emoji] theyíre getting there though [emoji] [emoji] [emoji] @ Bodyfreaks https://t.co/k5EPOpnxu9	0.445	0.054
· #SHINee # [emoji](@SHINee) 'I Want You' Dance Practice#SHINee_TheStoryofLight [emoji]#IWantYou #SHINee_IWantYou # [emoji] https://t.co/At2H9BG6We	0.444	0.053
· on behalf of me and all of @Argentina, weíd like to say that we love you guys and are with you today! vamos @NGSuperEagles!!	0.443	0.053
· I need this oh it's so cute!!! https://t.co/ANAtgG4vzV	0.443	0.053
· Oh @realDonaldTrump seems there are world leaders now making plans for your future. This is today's reminder that the Hague will be your moment of truth. And me I'll be having a BBQ and celebrating you rotting forever #NaziPrez #LittleDonnyTraitorWannabeDictator #IMPEACHTRUMPNOW https://t.co/QEWdPwt7r3	0.442	0.053
· Ahhh this girl, I love this girl v much! [emoji] @januege https://t.co/Cit1ZqoI05	0.441	0.053
· Omoooo I'm seeing two georgous men....But wait, georgous isn't enough to describe their beauty. Aaaaaahhh [emoji] [emoji]@BTS_twt #BTSxLotteFamilyConcert https://t.co/3uci26Wulk	0.44	0.052
· @XeniaKaepernick Thank you!! Iím loving it here so far [emoji] [emoji]	0.44	0.053

· IM SORRY BUT IT CANT STOP SPAMMING THIS I LOVE HER SO MUCH https://t.co/d3k7x4h4Ud https://t.co/d3k7x4h4Ud	0.44	0.052
· iím so overwhelmed. theyíre just beautiful :í) https://t.co/9bHwb5tode	0.439	0.052
· Hereís fun! 4 hours into our trip to Alton Water for the #greateastswim & weíre only just passing Crewe. The scenic route is now as bigger a carpark as the M6 weíve been to 2 McDonalds and our youngest has been sick. STILL STRONG [emoji] #ididitforthechristie #teamchristie #greatswim https://t.co/Wyi348jQDP	0.437	0.052
· happy birthday to my queen @sendatte [emoji] [emoji] hereís a lil throwback where weíre cute lol https://t.co/rUHxXVMJPY	0.435	0.051
· @spiritedlunakat Iím so happy I got to see him live [emoji]	0.434	0.051
· YES! I'd show up if I were anywhere near NYC #FoxisPRAVDA https://t.co/nxkwTT4CR2	0.433	0.051
· OH MY GOSH IM AO EXCITED [emoji] [emoji] [emoji]AHHH THIS IS AMAZING https://t.co/8YrDcsptB9	0.431	0.05
· Na so jare!!!We stay optimistic We stay supporting our #SuperEagles#GOtvAfrica#BringRussiaHome https://t.co/szkPb3FhJM	0.43	0.05
· Weíre #1! https://t.co/4wTi3KaqJh	0.43	0.05
· ONLY YOU #43 on iTunes in the US [emoji] [emoji] [emoji] [emoji] [emoji] [emoji] LETíS GOOO Itís getting up there @CheatCodesMusic @LittleMix https://t.co/89kbs7jHab	0.429	0.05
· I've just watched episode S03E20 of Blindspot! #blindspot #tvtime https://t.co/98T6wdoccO https://t.co/qqEIF4TJea	0.427	0.049
· Good morning piggies. Itís a beautiful to go into debt for me. Shower me with dollar signs #findom #FinancialDomination #paypig #humanATM @rtfindom @RTfaggot @SupportDommes https://t.co/zu62Drst5P	0.427	0.05
· Iím really not getting too excited yet. Still need a LB, RB, GK, Striker x2, CM.#NFFC	0.427	0.049
· God, I love him SO much. [emoji]#SaveShadowhunters #SaveTheShadowWorld https://t.co/chNIcFIy3G	0.426	0.049
· I CRY [emoji] He's so beautiful UWU [emoji]@BTS_twt | #BTSxLotteFamilyConcert https://t.co/iSvQcbDgrP	0.426	0.049
· Wishing @NGSuperEagles all the very best, let's play like its our last,play with everything we ve got, millions are on your neck counting on u to make us proud #GOSUPEREAGLES #SoarSuperEagles #NGAISL #WorldCup #Team9jaStrong #SuperEagles	0.425	0.049
· SHUT UP IM SO HAPPY EHAT THE FUCK I LOVE LARA JEAN AND KITTY HER SISTER OOP	0.425	0.049

· My husband thought this campaign video was a TV show. It's good. It's veerrrry good. Good luck in your campaign, @mjhegar ! If I lived in Texas, I would vote for ya https://t.co/Raky8yMBzS	0.425	0.049
· Weíre so excited to be the only Ottawa location carrying new hardcover editions of #Kagagi by @JayOdjick! We have signed copies for your #nativebook #graphicnovel #collection. Get yours, theyíre going fast! https://t.co/6dnCsoiwpm	0.425	0.049
· Wait, hold that last tweet! Iím gonna stream later, after @irunthemc streams bc heís doing that Octopath demo again. Apologies to electricalskateboard, slocool, & cutehealgirl. Iíll see you all around 11MTN	0.424	0.049
· I've just watched episode S08E09 of Desperate Housew...! #DH #tvtime https://t.co/ztfUNMPBRt https://t.co/PwS2fo35pS	0.424	0.049
· New products added on my store! Check them out now! They are really cool! - https://t.co/qiyR2OI6DX	0.423	0.048
· Next week, we're working with @kinsellatrust to celebrate Ben's life and #stopknifecrime. Stay tuned! https://t.co/wIL5npBvvq	0.423	0.049
· @yesnicksearcy I'm watching @JustifiedFX again. Quality Entertainment!	0.423	0.049
· You think youíre SPICEY? Weíll see - @madalchemy is presenting the SPICIEST RIDER award this year! https://t.co/OXRkwk8edX	0.423	0.049
· I liked a @YouTube video https://t.co/old7IUgdkL [emoji]WORST THROW-IN EVER! [emoji] (World Cup 2018 Ronaldo Iran Spain Uruguay Saudi Arabia	0.422	0.048
· 13. HYDE - female pirate ver. I've been in love with this costume for years and want to do It! It's such a beautiful costume [emoji] [emoji] https://t.co/byCSHPEdST	0.422	0.048
· We're #hiring! Read about our latest #job opening here: Deli/Bakery Clerk - E Markland Ave, Kokomo IN - https://t.co/RzNHkyw8B5 #CustomerService #Kokomo, IN #CareerArc	0.422	0.048
· Waiting for you to come im so hungry [emoji]	0.421	0.048
· OH MY GOODNE S I TH KM IM DEAD https://t.co/4Q950TxDtm	0.421	0.048
· iím glad this isnít true for me because iíd probably be dead by now https://t.co/3ssXozxzft	0.42	0.048
NEGATIVE: Dimension 4		
· @zazingwa @CryptoPseudonym @Crypto_Trogdor @cryptostardust @Elbrusco3 @mrdmrts @cripdohsimpson 4/ Can you explain why anyone would pay the astronomical price to fake the launch of hundreds of satellites, and hundreds of thousands of similar videos/images? What	-0.626	0.106

benefit is gained by such an extraordinary, and arbitrary, conspiracy?		
· @iamkhaledd No, this country wasn't founded by immigration... What are you talking about?The British came here and conquered the land, then maintained it.	-0.619	0.104
· @InuOfLegend "Finally actually giving your younger sibling attention?" Inu asked though she of course can't even trust what love is showed for her.	-0.596	0.096
· @TanaLara4 @Punkster1011 But all models do pose somewhere nude. What do you think PlayBoy Magazine is where almost most famous people pose nude? Isn't that "Adult" magazine?	-0.57	0.088
· @adjani_98 Right? How much common sense does it take understand how they can't make more money by helping less people have it....	-0.565	0.086
· honestly donít understand why people eat animal products when plant-based tastes better, is healthier, has less of an environmental impact and doesnít hurt animals. Kind of a no-brainer, donít you think?	-0.548	0.081
· @UnimpressedTX @SfaLumberjack21 Which in most cases is a joke. Online registration sometimes requires drivers license check, but in most states you don't need to prove citizenship or even legal residency to get one. The default is the "utility bill" which verifies nothing.	-0.54	0.079
· @queenyuckyuck @stavvers @transscribe "women do not have a prostate but have a gland that functions almost exactly the same", so there's thing, uh, Reading comprehension?	-0.534	0.077
· @christinawilkie What ever happened to the National Enquirer just being crackpot alien abduction stories, sasquatches and mutant bat babies? Used to be a respectable publication.	-0.534	0.077
· @lmao_hi1 ".............''He glance back to her with an innocent face. He is just as confused as her right now, he can't remember anything last night after he bring her to her and they have a drink together. "What... Just happent?"	-0.526	0.075
· @CBCNews Again .. why is the Canadian Broadcasting corp. taking even 1 minute to post this as any kind of issue .. considering this happend in another country ... is this the best you have. ???	-0.518	0.073
· @adrianowlett @SCrabbPembs @cristo_burton Well, a Norway plus style of Brexit has at least elements of pragmatism and sense. But maybe attacking someone for trying to respond to this disaster is not good strategy?	-0.517	0.072
· @CillizzaCNN Yes you are right. Since 1997 law enacted by William Jefferson Clinton. And in 2014, when it was learned that Obama was following that same law. Did you show your outrage then? Or maybe, like everything else, not much coverage, because Obama was your guy. Complicit or what	-0.512	0.071

· @PetterSolbergn Flynn admitted under oath that he lied. Was he lying when he said that? He wasn't railroaded. He admitted guilt voluntarily in open court. You still have no answer. If the meetings with Russian Intel were nothing, why did they lie? KushnerSessions, Stone & Prince lied under oath	-0.509	0.07
· @LukeGromen Mark just trying to face the sad fact that UK doesn't have a piss pot amount of gold left. Sold at $250. Lots of diversity though.	-0.502	0.068
· @masonasons @Shaysters So what happens when you shit out the poop tart? What do you call that, poop tart squared?	-0.501	0.068
· @CBCAlerts Why didnít Canadians detain the hundreds of Americans who got blown over to Canada last year in Port Huron during the Port Huron Float down and didnít have any form of ID?	-0.5	0.068
· @HarveyLevinTMZ whats the story line.. hubby divorces her for what? or did she do the passing thing?	-0.495	0.066
· @realDonaldTrump When are you going to start making your stuff in the states? Are you going to have tariffs on the items from your companies coming in from overseas? Just a fun fact for you, you are only one point away from Nixon's impeachment numbers.	-0.494	0.066
· @CancerEdInst @FDATobacco What's the danger of flavors? Vaping is not a gateway to smoking, as smoking rates keep falling. How do these "flavor dangers" compare to smoking? Why are you trying to protect smoking and cigarettes by denying basic human rights to adults??	-0.488	0.065
· Can anyone explain why people convicted of felonies lose the right to vote? It doesn't make sense. https://t.co/e9CbC6iShC	-0.484	0.063
· Simple technique to find out whether the other person can be trusted ?Observe the behaviour of other person towards you when you, the other person and the third person are having a conversation.Do this for few months and you will have a answer in YES or NO.	-0.483	0.063
· @MichaelAvenatti I wouldnít exactly say ìyoungî. Plus, how do you know where they are going and/or why? You are an ambulance chaser and nobody has any requirement to tell you anything. Sounds like you are probably making this shit up like the rest of your ìcareerî	-0.483	0.063
· Islam didn't make it haram to fall in love. It didn't forbid you from wanting someone. It only guides that love so it protects you, her, your families and especially saves you from humiliation on judgement day.If you love her then why are you so ok with letting her engage?	-0.482	0.063
· @Ptmurf1016 Hahaha agreed. Trying to make sure the conversation doesn't stay 100% single tracked on immigration while the robbery continues.	-0.481	0.063

· @freedom_moates No different than needing a warrant for anything else. Who says Law enforcement isn't put in check.	-0.476	0.061
· @SkyNewsAust @annelisenews how many times has someone bumped into you wearing headphones. you need to be focused on your total environment to protect yourself and not immersed in what is coming thru the phones. its called spatial awareness and could save your life.	-0.475	0.061
· @charliekirk11 Jesus? Your comment explains a lot about you and the small minded misguided anger you posess. You my friend were never saved.	-0.472	0.06
· @SpeakerRyan addiction is a medical condition what is your plan to treat the condition and to not punish the afflicted? #addictionisadisease it is not just people taking opioids. #ResignNow https://t.co/kHE0UhPDnV	-0.466	0.059
· 3. Either the policy OR the means of making the policy might violate the law. For instance, the policy violates the ADA by assigning people to the wrong care, or the policy was made in violation of state good government laws, or substance or procedure violates the const (4/?)	-0.461	0.058
· @APICGLAC Question: Any literature or thoughts on Nursing staff dampening towels (washcloths) then microwaving them so that patients can wipe their hands with them prior to eating their meals? Similar to what the airline industry does for their passengers.	-0.46	0.057
· @EdKrassen "The native Americans were wiped out when foreign peoples invaded, so you should let foreign peoples invade you." What could go wrong?	-0.46	0.057
· When caught doing something wrong, or trying to hide something wrong, Trump always falsely accuses the other side of doing the exact same thing. Trump, before cameras, accused Democrats of running a ìchild smuggling ring.î Is this why thousands of immigrant kids are ìmissingî?î	-0.458	0.057
· @AliaChughtai @AJEnglish @AJENews Can you please take a screen shot? Site is blocked in Bahrain.	-0.456	0.056
· @IsabelOakeshott To avoid the trolling (although this is what good journalists do until they get their story) why don't you just release what you have and admit brexit is stupid idea funded by Russians ?	-0.455	0.056
· @CillizzaCNN Her super sly way of telling the Mockingbird media to stuff it, was priceless. How priceless you ask? You shitbag's can't stop talking bout it. Sheer brilliance. You idiots fall for the trap everytime. His approval goes up, your credibility goes down. #WWG1WGA	-0.455	0.056
· @celeste821 It was a yes or no question. Is it okay to call someone 'evil' because of their ethnicity? Granted there's no point in asking the question, since you've	-0.454	0.056

already established that you think white = evil, much like how Hitler established Jews = evil.		
· @SteveHusker @pigchampions @biglargeben @realDonaldTrump why do you think he was crucified? he broke the law...	-0.453	0.056
· @shamus_clancy Not sure that Kawhi is the mentoring type. Doesn't appear to be based on his demeanor but you never know.	-0.452	0.055
· @Zupaku Giving her hair a light ruffle Tetsu removed his hand, smiling at the sight of the flower."Are you gathering flowers?""But where are your parents, are you alone?"	-0.451	0.055
· @AUThackeray @UNEnvironment @ndtv @dna @deespeak @timesofindia how come media and celebrities not promoting or endorsing these kind of good initiatives?	-0.448	0.054
· Did you know that you're probably breathing incorrectly during exercises? For efficient breathing, breathe out when doing activity that curves the spine. Breathe in when your arms or legs move away from your center, or during any activity that straightens your spine.	-0.443	0.053
· @ANI Statistics have a unique quality of reflecting only what you want to show. There's always another side of the coin. Why not quote the figures where the defense budget was rejigged or the influx in soldier's death.	-0.442	0.053
· @PMOIndia is it possible that a person having different name in all the documents for passport can get the passport if he/she approaches the media.	-0.441	0.053
· @danmericaCNN Why does anyone feign surprise at this? TRUMP IS AND HAS ALWAYS BEEN A CONMAN AND EVERYBODY KNOWS IT! He has no more idea about how to do the job of President than a 6 year old. In fact, a 6 year old could do better, avoiding tearing at the fabric of the society.	-0.441	0.053
· @brian_armstrong @coinbase @yin_amy Nice buddy why donít you hire some women to do support?	-0.44	0.052
· @realDonaldTrump You obviously don't even understand Trade 101, do you? Less tweeting - more learning the basics!	-0.439	0.052
· @Sanele_so We have all been hurt my friend but not everybody will fuck you over and sometimes we hurt people even when we donít realize it. It could be a silly comment or whatever the case may be and that is why forgiveness is significant in our lives.	-0.439	0.052
· @JoeConchaTV What a douchey thing to say. Telling people who voted for trump that they are at the border like Nazis and picking and choosing. What the hell are you talking about?	-0.438	0.052
· @rainwater_ip @rachel_maria65 @SoitenlyBob @realDonaldTrump And how much did the Hillary camp	-0.438	0.052

spend on the fake dossier? Seems like there was a hell of a lot more collusion going on with her camp than with Trump. Nobody wants to talk about that though		
· @freedarko ah yes...the fascism of the left again rears its ugly head.Hey pal, you know people can freely support and or vote for whoever the like in America. Just because you disagree does not mean you can try to destroy their career and lives.	-0.437	0.052
· You aren't born in a lotus. Dalit or Upper, the life is sexually transmitted in a worst way possible. Burn that supremacy....	-0.436	0.052
· @EPA Confidence ? F that .. how about actual science driven public data driven policies .. Pruitt is an environmental terrorist making tons of dark money	-0.434	0.051
· a lack of empathy hinders solutioning humanity issues in hypotheticals (imagination), even when faced with empirical data. a lack of empathy hinders acting on humanity issues in the present. a lack of empathy will not relate to experiences other than their own	-0.432	0.051
· @simonrug @ccurts2 @SteveinSenni @davesparade Do you think the games could be moved around local prem club grounds in the area? This set up is very positive and should help continue to develop players and strength	-0.431	0.05
· @nilbogmilkk @genet1245 Do you know what words mean? Specifically, all of them	-0.43	0.05
· @Arron_banks Says the man trying to score political points - while ignoring the fact that it will mean thousands of job losses. None of which will affect you of course, so you know, tough luck because ësovereignty and blue passportsí.	-0.429	0.05
· @kylegriffin1 Wow, Fox makes the argument that as long as children are not Americans it does not matter what happens to them. I hope all of the Fox News sponsors take in the horrific opinions being expressed on Fox. It matters what happens to every single child.Shame on Fox.	-0.428	0.05
· @LateNightSeth have we acknowledged that trumps signing executive orders with a black marker and not a pen. Who does he think he is, signing playbills at the stage door?	-0.426	0.049
· @BeardedGenius Such a sin the baby is crying because the mother doesn't give a damn about her.	-0.425	0.049
· @POTUS @therealdonald_ @RealDonad_Trump DNA testing of all incoming illegals is necessary to ensure accompanying adults with the children is not a human trafficker. This DNA can be used to identify crimals.	-0.425	0.049
· @Raffi_RC @POTUS Trump and Trump supporters who claim to be christian are what the bible reffers to as "Luke warm Christians" Trump and company, however, are a bit worse.	-0.425	0.049

· @kiwimikewinton @jllgraham @MikePenceVP @realDonaldTrump What do you think the flag represents?	-0.424	0.049
· @Duece1324 rankings don't measure heart, dedication, determination, and the support system you have. https://t.co/ocRXWNmh1G	-0.423	0.048
· @Peston @sajidjavid Clueless the fkin lot of them.... what a waste of an ëeducationí.	-0.423	0.048
· @madhukishwar @PMOIndia @SushmaSwaraj Just to cajole Islamists, there r plenty of Hindus 2be made Scape-Goat! What wrong Vikas Mishra has done? Is it sin2 ascertain the credentials of a Pass-port Applicant? A LIU personeel comes2 Pass-Port Applicants' residence & find everthing OK,But still begs for bribe,is it OK?	-0.423	0.048
· @SpeakerRyan By cutting Medicare and Medicaid you show that you have no heart. (For those who still havenít caught on after your tax scam for the filthy rich...)	-0.422	0.048
· @rjward86 Have you checked the aerial connection is secure? Are you able to check this on another tv may be that has freeview built in? ^Dee	-0.422	0.048
· @defend_bc You are demanding someone else pay for it. You don't have a right to the property of others.	-0.42	0.048
· @realDonaldTrump Poor baby! Got a BS in economics, doesn't understand economics. Did daddy pay for your degree?	-0.419	0.048
· @swanmoonbaby @ACLU You know that little girl in your profile pic? Her mom ripped that girl away from father without notice. And that girl was never separated from the mother. Youíre another victim of the fake new	-0.419	0.048
· @Joel134748 @SyeTenB By all means, jump in the ring and handle business, big guy! Donít just start calling names, start throwing punches!	-0.418	0.047
· @AAC0519 @nsewanswers @nytopinion This is a lie. The zero tolerance and separation of families by the thousands has not occurred under any previous administration. Not even Trump disputes this.	-0.418	0.047
· @MeghanMcCain Please do not waste your energy criticizing the Democrats on The View. Please do something to help bring back the true American values that your father believes in. The Republican Party will be ruined for good if nobody does anything to oppose Trumpís evil deeds. You can do it!	-0.416	0.047
· @Tucker35lre @Nadeshot @FortniteGame And only listening to casuals does what? Kill games. Look at h1 just bc people couldn't aim they messed up the aim system. There's an old saying... Get good..	-0.416	0.047
· @JamboEveryone @_kimhanley @DisCounselor Designed by Joe Rhode, a Haunted Forest would be a great addition. How about a full AA Goonies ride ?	-0.416	0.047
· @GOP @TeamTrump 1.5 years of witch hunt and ZERO EVIDENCE of Russian collusion. DNC server never	-0.415	0.047

investigated by FBI. Fake pee pee document paid for by Clinton. Christopher Steele terminated by the FBI for lying. FISA Court judges deceived. Conspiracy or incompetence?		
· Does it strike anyone else as odd that the official #worldcup ball is ìTelstarî? Wouldnít ìSputnikî have been a little more apropos of a Russian tournament?	-0.412	0.046
· @JTaylor_187 JT, you joking mate? When has our judgement ever been clouded with blind optimism??	-0.411	0.046
· @KurtSchlichter Another anti Trump Republican that refuses to accept that people do not care what he thinks about the POTUS as evidenced by his recent loss at the polls in South Carolina.	-0.41	0.046
· @NahazDota @breakycpk ToS: may include functionality designed to identify software or hardware processes or functionality that may give a player an unfair competitive advantage when playing multiplayer versions of any Content and Services or modifications of Content and Services (ìCheatsî).	-0.409	0.045
· @pieter44408533 I never said it was. I canít understand how the word white, which is portrayed as a color of purity and positivity offensive and racist. Especially when no one black made that up.	-0.409	0.045
· @renegoac1978 @TorontoComms Thank you, this has been forwarded to the Street Furniture section for clean up. ^de	-0.408	0.045
· @Holidayer16 Hi Jason; were you able to download the voucher when it was first sent through to you? GM	-0.408	0.045
· @Super6 who were the last winners of the £1000 from the last Jeff head to head?	-0.406	0.045
· @HaltonPolice @DaxMc The location identified is a school crossing guard location. Pedestrians do not have the right-of-way at this type of crossing unless a crossing guard is present and displaying a stop sign. https://t.co/JcZ4nhhysq	-0.405	0.044
· @duffer007 Hi John, do you have a dial tone on your home phone? Em	-0.403	0.044
· @sadierodgers8 @mamfe71 @skankyliar @warningshout @Fifif75 @TruthTe113r @ILoveaDiddyman @KatiePrice There will be a very scripted conversation of the aftermath, one where everyone gets their stories right for the cameras. Let how many robbers there were , 3,6,9,12????	-0.403	0.044
· @realDonaldTrump What does that even mean? More nonsensical jibberish blaming Democrates for your failed leadership. You need Dems because so many Republicans oppose you. By the way, are you aware that Rebulicans control the entire government right now. So clueless.	-0.403	0.044
· @Tilly_brukz Why did you move his birthday to between the last match and today? Lol	-0.402	0.044

· @mangobaaz @Doppaminee @CreepoPotato @mangobaaz why isnt the link not opening?	-0.402	0.044
· @MinaMorcos_ @KyrillosSaid Alright somebody killed someone else. Now you kill the first person but then you gotta get killed because you killed the first person. Then when you get killed somebody needs to kill the person that killed you and so on. How does that fix anything?	-0.402	0.044
· Self-Efficacy is defined as an individualís belief in his or her innate ability to achieve goals. Albert Bandura defines it as a personal judgement of "how well one can execute courses of action required to deal... https://t.co/OSprZSzT07	-0.402	0.044
· @FlatEarthOrg ya know right that the greeks knew that the earth was round long before nasa was a thing right? And nasa is not the only one that arrived in space and got on the moon	-0.402	0.044
· @SarahPalinUSA @northsooman Total liberal Idiot. This cannot be defended. Only reason he is on Fox is for Fair and Balanced. Hell he got fired from a liberal news network years ago. Even they thought he was over the top	-0.401	0.044
· @AdamBlease @sampower_sam @showcaller @sebshawbarnsley @m1ckeyjoe Poor Mr Shaw. Does that theatre not have a back door he could have escaped through?	-0.4	0.043
· @fred_burton @LuluLemew @CNNPolitics Why is Trump redirecting US Marshallís, National Guard & JAG lawyers to work immigration? They have their own jobs to do. Military bases a kiddie jails. This administration is beyond FUBAR.	-0.4	0.043
· @PiyushGoyal IPO may get over subscription but it doesn't multiply in numbersHAL BDLetc etc Ppl still crying for original price they paid while subscribing 6 months backGov gets money from ppl but the fact is that get trapped	-0.399	0.043
· @biffblisters @kimmyifuplease @elee1025 @warrior_4_good @AVestige1 @ExpectoResister @PurpleDahling @milkexperiment @Zylie @TheUSASingers @Distracted66 @islandertmt @SelfImposedXile @McnicolSalazar @LakeCountry @bjcreigh @Chowder_Society @fmc23169 @Coldhands2 @Peterandfam @Pacoluismonta9a @Jabbadaddy2016 @MissGFYCuffy @j4hub @BlazedSadElle @DanteUSAInferno @nullnotvoid @CannabizLawyr @Jim26128472 @RutherfordRocks @deb7519 @Marion_aruaL @eronel35 @WhyNot_RESISTS @ACJJustice @Orthotottie @Bellarealness @myworldjlt @FineYoungAnimal @Nottoofondofyou @HotWifeyHeather @TanousLisa @Lou_Duderino @BarleeBlue @pebbles9010 @DumbleDarkVader	-0.397	0.043

@Citizen_13_ @DarkDNTM @BmoreTrell @mysoftsofa That man has a nice firm butt. Anyone have a quarter to bounce on it?		
· @NeolithicFarmer @Sheppard250 @RichardBentall @LeeSeater @jaminbjoel @donnyc1975 @richiebee Really? Is it that difficult & complex for you to figure out????..... maybe, just maybe, therein lies the answer to your original question....	-0.397	0.043

POSITIVE: Dimension 5	coord	ctr
· ARE YOU HUMAN TOO? [emoji] Where did you find that extra charm? You're killing me softly! WHY SO HOT?? PLEASE EXPLAIN [emoji] [emoji] @BTS_twt @bts_bighit #BTSxLotteFamilyConcert https://t.co/oUhE8dkYZy	0.702	0.156
· Find out how much you could save switching to BrandStencil for your artwork creation - use our ROI calculator! https://t.co/vro5K44Q2n #MarketingAutomation	0.644	0.131
· Attn, #mixers! Stop what you're doing and listen to @LittleMix and @CheatCodesMusic's #OnlyYou [emoji] [emoji] [emoji]†#littlemixONLYYOU https://t.co/AjkcT0zWqV	0.643	0.13
· All #robots need drive systems to move around but do you know how it works? Join our #STEM courses for free! https://t.co/RCmUjar9kX	0.641	0.129
· Where are our #DogsofHOS at? Donít forget to tag us in your photos so we can get over-excited and share them everywhere! [emoji] https://t.co/tBZEzn1tZn	0.627	0.124
· @PPActionCA When will forward- thinking California learn what Connecticut already knows? #Essure is not safe or effective! #essureproblems https://t.co/Rrsv6iydn0	0.61	0.117
· One of the first things to learn about this journey is that itíll be challenging. Your prospects may have a million reasons upon missing your appointments. However, the challenge relies on your reaction. Watch on VTube+ now! >> https://t.co/UJY8h3C2ao #TheVLimitless https://t.co/UmZD9eKRVL	0.602	0.114
· Here's your @AFCU [emoji] official Fan Favorite T-shirt of #TOU2018! Thank you to everyone who voted! Be sure to pick one up at the expo during the race. [emoji] https://t.co/NtXHgq5rf1	0.594	0.111
· @SoonerStark Yeah. The worst are these:Politician: we are happy to announce weíve cured cancer.Commenter: well what about children dying in	0.589	0.109

Syria you POS!!?? I hope you die of a mutated form of cancer and your whole family!!!		
· @TomArnold Lol haha your so stupid! Didnít you know if anything exists MuleíR has it. And guess what that means when evidence is sealed? Youíre going to fall... you believed him [emoji]	0.588	0.109
· How to design effective website navigation? How to decide what's important, avoid common pitfalls & deal with navigation in SaaS and eCommerce. Listen to this UI Breakfast Podcast with @els_aerts https://t.co/7m1KUjq8fN https://t.co/grHnEmJSPC	0.572	0.103
· Hey, I have opened my own online store. Do check it out and let me know what you think? https://t.co/6EeRhARyPh	0.565	0.101
· Want to get a new car, but arenít sure how much your trade-in is worth? Find out here!https://t.co/eS6SjGAocG #TradeInValue #CarBuyer #CarValue	0.563	0.1
· COME ON CHARLOTTE SHOW UP AND OUT! #getUsome #brunchSoHard #gummyblayr https://t.co/3rf1buecVO [emoji] [emoji] [emoji] https://t.co/c4HrFwAleR ó feeling fabulous	0.546	0.094
· Just joining in on the fun? Don't be shy! Remember to RSVP to our #benandhollyparty to be eligible to win these prizes! AD https://t.co/asEhBmM0sb	0.545	0.094
· is bts must suppose to be bangtan sonyeondan? no. did they wrote BTS just to gain likes from manh army? no. BTS isn't always bangtan sonyeondan. it can be behind the scene too. then do you ever think that way? no. think rational. open ur mind and don't be stupid. #BTS #ARMY https://t.co/wbAMFxyDus	0.537	0.091
· @realDonaldTrump Outch !!!Can U say "Numbnuts" !?!Hey The Mooch !!!Miss your old pal ?!?Cheers !!!Ern [emoji]@Scaramucci @FLOTUS @POTUS @CNN @cnnbrk @colbertlateshow @StephenAtHome #MAGA2018 #TrumpConcentrationCamps #melaniajacket #Trump #TrumpLies https://t.co/G5PqZGtIul	0.528	0.088
· Art [emoji]Shanghai EXPO 2010Back in 2010, China hosted the #ShanghaiEXPO.We landed in Pudong, #Shanghai at night and did I say 2010? It felt like 2024! It was nothing like I imagined. Read our blog on tumblr: https://t.co/CNEXNqHRSc#artworld #tbt #inezsuen #china #worldsfair https://t.co/QfZS5DdEI6	0.526	0.087
· Simple technique to find out whether the other person can be trusted ?Observe the behaviour of other person towards you when you, the other person and the third person are having a conversation.Do this for few months and you will have a answer in YES or NO.	0.526	0.087

· @dcromp84 @realDonaldTrump You are correct! Obviously as a former policeman you understand that President Trump is enforcing and strengthening the law to keep all American safe! For love of country and citizens first! Preach on brother! #MAGA	0.524	0.087
· Love this definition of ROGUE:R-relevantO-organizedG-groupU-undergroundE-educatorsAre you going ROGUE with your PD? #40CF @4OClockFaculty	0.52	0.085
· @Murphys0311 @SeasideSunsongs Oh, man. Iím so sorry to hear that! I hope your healing goes smoothly. Hit me up for puppies anytime.	0.518	0.084
· @sabrinaflies Stop what you're doing and listen to 98.1 KDD! NEVER BE THE SAME by @CAMILA_CABELLO is about to play! https://t.co/IFO7UEbwVC	0.517	0.084
· It's FRIDAY! Map Girl!! You won't believe what states we sold vehicles in this week, and listen to some CRAZY facts that I bet you didn't even know! #mapgirl #wowwoodys #needacar #carshopping #kansascitydealeship #funfacts #customertestimonial https://t.co/rXWJsLivgt https://t.co/jTy0E0wEfe	0.514	0.083
· @AmbassadorRice You know that this goes on at all border areas, right? I mean you were an ambassador for heaven's sakes! Drive down the Northway in Upstate NY or leave El Paso airport. Border patrol has a job to do.	0.513	0.083
· Summer is here and the sun is out! Donít forget it can be very helpful to have people who can support you when you are having difficulties, no matter the time of year. https://t.co/YuG0vB3mE2	0.511	0.082
· What if the cast of #LordOfTheFlies were all women? An interesting exam question? Book your school in now to see our co-production with @ShermanTheatre and get a FREE interactive workshop! https://t.co/aIEDNBCYIX #englishgcse @wjec_cbac @CCEA @Edexcel https://t.co/FtuEQNFHgj	0.509	0.082
· @5SOSFamMsia Dont worry u wont be lonely#YOUNGBLOODListeningPartyOnHitz	0.509	0.082
· you donít know what itís like !!!!!!!	0.508	0.081
· @GOPChairwoman Call the dems PROGRESSIVES.They want no borders and lawlessness like Europe.Where are you?!?Get out your megaphone please!Collate clips of all the progressives spewing their anti-Amerecan rhetoric.They are leftists now.NOT liberals.America needs you, Ronna.	0.508	0.081
· @stonemirror @realDonAnon @Metascover @John5x5 @Avsniper @fortiveli @MudRemover @fasteddyTO @Gzonnini @stackchipsdaily @TrumpisintheWH @Muoloc @Ellen4Trump @Quantum_Stoic @skr52562 @NowellKern @luvleebutterfly @Chilon_Sparta @jleadsufollow	0.506	0.081

@Duceman03 @jokinandtokin @patientsan @PissedOffPatri3 @RightDaisy @John81726765 @HavokHawk @garos56 @Veritas_2016 @PetterSolbergn @Texas_Eric1 @TheWayOfMatt @HouseofPain1776 @ArchieWouldSay @wildlillie @ahnonymous_a @capron_bruce @BeachhouseBabe @Silvana5933 @qanon76 @BeckiDoll007 @MagniFieri @Thomas1774Paine https://t.co/kgC8ondzJD What! It Canít be true! Lol #MilitaryTribunals #WarCrimes #RICOAct #Treason #EnemiesOfTheState		
· @kylegriffin1 WTF?? Seriously? ì Papers. Give me your papers?î Itís a scene from Nazi Germany. Im a US citizen. I donít carry my passport everywhere! How many carry their passport around on a daily basis? Just because I have a state ID to drive doesnít mean Iím a citizen.	0.506	0.081
· @ahfyujin Yeaaaah! [emoji] x D what are you doing right now, Yujin-ie?	0.505	0.08
· @PressSec @WhiteHouse How else do you get liars to tell the truth? You've ignored everything else, trying shame to see if that has any effect!	0.504	0.08
· @kenrentz @NBCNews Your kidding right?You don't know what the WTO is do you?You just posted a smartass comment to sound smart? Butt, you didn't!	0.503	0.08
· @CillizzaCNN Yes you are right. Since 1997 law enacted by William Jefferson Clinton. And in 2014, when it was learned that Obama was following that same law. Did you show your outrage then? Or maybe, like everything else, not much coverage, because Obama was your guy. Complicit or what	0.502	0.079
· @WindsorHugo @LeoFrielPhoto @Steviegrieve @Scotpol1314 @RedBrickz26 @ScotsSolomon Feel free to supply evidence that refutes my figure...Scottish employees having to uproot and move South to find work?Are you having a laugh?That's been the case for over 30 years pal...that's what happens when the UK economy is centralised in and around London.	0.5	0.079
· @PressSec Seriously? Try as you may, it's about time you realize that nobody looks to you for an ounce of credibility, you lying sack of crap!"it's the Dems", "it's the Dems", "it's the Dems", "it's the Dems", "it's the Dems", "it's the Dems", "it's the Dems", "it's the Dems", ...	0.499	0.079
· @mainameiskai SZN is coming sooner than y'all think. Don't say that I didn't warn ya!	0.493	0.077
· @ddgmina dont talk to me, im sad. jk. how are you today, love? anything extravagant happening today?	0.49	0.076
· Interested in being an #RChain validator? #RCON3 will help you gain the knowledge and skills necessary to contribute to the #ProofOfStake validation process. Contribute to decisions related to fee structures!Register -	0.489	0.075

-> https://t.co/KwNJWH1zoRRSVP --> https://t.co/myozMKUpLw https://t.co/zBX8X3zWBk		
· @MollyJongFast @PressSec @Pontifex How much longer are you going to allow yourself to be an easily manipulated sheep, Molly? Are you getting tired of being a useful idiot? Wake up!	0.488	0.075
· @RichiTwoshoes @DavidWilletts3 @HandlebarWisdom @DefenceHQ Ok - so what evidence do you have that proves this to be incorrect ?	0.488	0.075
· BlackBerry KEY2 pre-orders start on June 29 and will be available to purchase on July 13Interested in your next mobile? Chat to our Comms Team #techteam https://t.co/lWpnTPwt7u	0.482	0.073
· Bless Our City plans are due THIS SUNDAY. Talk to your small group and decide how you can be the hands and feet of Jesus to our city with $500!	0.481	0.073
· @BillHemmer I watch you show this morning and the guy that was on ticked me off when he out and out lied . Saying that the right is calling people names and making fun of children in cages! Thatís not the right thatís the liberals!! Theyíre the ones name calling and swearing. Get it right!!	0.479	0.072
· Did you know that you're probably breathing incorrectly during exercises? For efficient breathing, breathe out when doing activity that curves the spine. Breathe in when your arms or legs move away from your center, or during any activity that straightens your spine.	0.479	0.072
· Hey #Friday....itís good to see you!! New ìBIG TIPî Dad Hat coming soon!! #LiveDreamBe #Digmi #TipYourCap #DigYourDream https://t.co/bZ2ODNkyUU	0.475	0.071
· Hey everyone, you know it & I know it. @comcast dish spectrum SUCKS!!Stop #overpaying for #crappy #cabletv! Worldwide [emoji] on any device take it with you [emoji] [emoji]Just go https://t.co/3qZak9I0XE!! https://t.co/kaZs4O4ykH	0.475	0.071
· So excited to be part of this program! Have questions about #Lightning or making the move to #Lightning? If so, let me know! Happy to help in any way possible! #LightningNow #IHeartLightning https://t.co/iUElO0CPBI	0.474	0.071
· Anime?Angel?WHAT ARE YOU KIM TAEHYUNG?! [emoji] [emoji] [emoji]@BTS_twt https://t.co/y0p4s0jcpT	0.474	0.071
· OMG! You have to see this. #BIGOLIVE. https://t.co/Da7w7ZG8Ld	0.472	0.07
· OMG! You have to see this. #BIGOLIVE. https://t.co/DcA0iSX4ke	0.472	0.07
· OMG! You have to see this. #BIGOLIVE. https://t.co/znpIJnwXmU https://t.co/8lEux11JG3	0.472	0.07

· @ArmaanMalik22 Aww!!! Aaru why are you sooo cutee?? [emoji] [emoji]	0.472	0.07
· Oh wow! How good a sport is #PaulMcCartney? Love it! #beatles https://t.co/vKQUsBlTP6	0.47	0.07
· Symantec Partners do you have questions? @SYMCPartners' newly formed Partner Service team is ready to help. Learn more: https://t.co/ocb1qP1Cx2 https://t.co/xt7OlMjwUW	0.469	0.069
· Yíall taking it far now. We all have our opinions and some of us may dislike his wife. But that doesnít give any one the right to call her derogatory names like those. THAT IS DISGUSTING!! Stop with the insults. Youíre crossing the line of freedom of speech. #SiyahBeyazAsk	0.468	0.069
· @RealDonaldTrump endorse whoever you wantó2018 is our year, @fladems will come together & bring change to FL! #TeamLevine [emoji] https://t.co/cH8dXwmU6A	0.467	0.069
· Dont wait for me to die for you to realize what youíve done/lost	0.466	0.068
· @LastPassHelp when will @lastpass form fills be able to populate textareas and HTML5 fields like number inputs? doesn't work in FF or Chrome for OSX	0.465	0.068
· @Iam__HumanBieng Welcome back Rahul!! hope ull be fine!!	0.463	0.068
· R on loop and i suppose what set it apart for me is how it doesnt...? sound like an anisong??? or a song where you listen and think "ah. is this from a game?"R sounds Out There, like a proper single song in its genre and i love that.also love how the guitars are deeeeep	0.463	0.068
· @realDonaldTrump Aaaaand... LIE! Dear God, your fingers touch the phone screen and lies come out flowing.	0.461	0.067
· Be Aware = BEWAREMonitor your blood pressure symptoms, save lives!!https://t.co/59SvCKvzxd#healthmonitor #healthtechrocks #winninwithhinnen #sahmsuccess https://t.co/HsDfkSkeVG	0.46	0.067
· RT if youíve ever heard of ìINTERNATIONAL TV AND FILM STARî Nick Searcy!#MAGA people are traitors. https://t.co/Idk2rqQBQJ	0.459	0.066
· @happilyerrin THANK YOU!!! LOVE UUUU	0.458	0.066
· Don't miss it you will regret later! https://t.co/pNoLo9ir8H	0.457	0.066
· @Jordan_miggy LOL I am so dead! How could you do this?! [emoji]	0.455	0.065
· Stuffed animals are back at Harley-Davidson of Madison! Stop in to our MotorClothesÆ department & check out all of these super cute fluffy animals - You're sure to want to take one with you!And don't forget the cute digs for the kiddos this summer!!	0.455	0.065

· @realDonaldTrump See hereís the thing, Mr Illegitimate, tweeting a lie doesnít make the truth untrue. Normal people know YOU are responsible for torturing little kids and that youíve done nothing to reunite them. You look very, very bad. #NarcissisticPersonalityDisorder	0.455	0.065
· @Mashiya46339260 @EFFSouthAfrica @Julius_S_Malema @Our_DA @COPE_SA Dear, you are not educated, you are disadvantaged, you don't know what supreme whites think.Keep voting for your party that takes all your money, their leaders live in obscene luxury on the money meant for YOU... You are too stupid to see that... that is why whites are supreme.	0.454	0.065
· Tired of renting and ready to own your own home? Broker, Kemar Johnson with McGary & Associates has many home buyer programs and can help you own your OWN home today! Call 954-830-2580 or visit https://t.co/Q18RnQMBTt to start the process or to check out homes now!	0.451	0.064
· @so_influential Oh no! That is not how we want you to feel. What's going on? I want to help. -Tiff	0.45	0.064
· @michellebb10 It's heartbreaking to hear the Mom tell him how much she loves her son. He's her love. Amor! We'll all play for our vile behavior. Stop giving money to countries that produce drugs, making wars aboard & deal w/problems here. People don't want to leave their homelands for US.	0.45	0.064
· Fridays always has me like this...don't know why! [emoji] [emoji]LOL https://t.co/dpedQX7vho	0.449	0.064
· ON SALE NOW [emoji] Tickets to @NightValeRadio at @TexasTheatre March 3rd 2019! Don't wait until the last minute get your tickets here: https://t.co/UptiSLHOiE https://t.co/UTbVzaJaHb	0.448	0.063
· @FatLandlord Take it you've never flown BA then!	0.447	0.063
· 4AD is taking over @sohoradio from 7pm! Tune in on https://t.co/TLuTmk6gmS #j‰gersoho https://t.co/bmGmJh0ADk	0.446	0.063
· Please don't forget to vote for # [emoji] # [emoji] [emoji] in Idol Champ today! https://t.co/S4dUNPuXgY	0.446	0.063
· After 56" slamming its almost-Have u taken action on R Madhav?Have u taken action Modi for visit?Have u taken action on ISI spy....?etcby @priyankac19 @IndiaTodaywith @PadmajaJoshi Do watch to understand how BJP speaking rubbish on J&K and their dumb policy. https://t.co/kK29oiYdQd	0.446	0.063
· @_EVANGELO PLEASE! They want to know what it is about you that makes you an asset to the company!	0.445	0.063
· Hey peeps hope you're enjoying your #Friday? Join us tonight for live #comedy & #music from @Laughter_house 8pm & #troublewitheva 10pm	0.443	0.062

[emoji] [emoji] [emoji] [emoji] [emoji] https://t.co/QOaU92cwet		
· Esos escalofrÌos que te dan cuando ves y escuchas arte. Pues eso es lo que me pasa con Kim Taehyung. Es puro ARTE! Those chills that you get when you see and listen to art. Well that's what happens to me with Kim Taehyung. He's pure ART! [emoji] [emoji] @BTS_twt #BTSxLotteFamilyConcert https://t.co/8TqcX1UulT	0.442	0.062
· ìYou havenít seen ëKangaroo Jackí?î Apparently thatís where @ColinHay influenced the young generation of today! #whatcoworkerssay	0.442	0.062
· Do you know who is awesome? My 1 new follower in the last week! Growing with https://t.co/BSuMCXJ6pq	0.441	0.061
· @btykiwi ill spam u dont worry ! and its ok letís just support our faves tho thatís why weíre here but yeah i feel u ìsometimesî [emoji]	0.44	0.061
· Do you remember when you joined Twitter? I do! #MyTwitterAnniversary https://t.co/AAA65KFVCi	0.44	0.061
· That's right. For all the "victims" and "poor families" the #Democrats say they represent? Look who #NancyPelosi met with instead? THE #ONEPERCENT! The #OnePercent billionaires club! To discuss their #ImmigrationReform Bill, not #AngelParents #KidsInCages!https://t.co/2Ckg4lk8lk https://t.co/36I97fLjcj	0.439	0.061
· @Astrid_NV @repdinatitus Glad to see someone is standing up to @FLOTUS ó I Care Melania, Why Donít U?	0.437	0.06
· Shout out to all the other lazy bums that do everything in their power not to get out of bed when they feel like their bladders are ready to burst. May you find that position to lay that makes it the least uncomfortable!	0.433	0.059
· Retweeted BOTSWANA [emoji] [emoji] (@EMCEE_LUX):Stop announcing you are single everyday!!we've seen it!"We don't want YOU" [emoji] [emoji] [emoji] [emoji]	0.432	0.059
· @foofighters There is hope, inbound plane has arrived in #Tegel Might just see you yet! Please #britishairways fly fast!	0.431	0.059
· IRR in USA: ëSA not high-profile issue in US Ö which is why it is so important that when it gets on the Washington agenda it is for the right reasons ñ not EWC.í IRR CEO Frans Cronje, in Washington. Take a stand with the IRR at https://t.co/8XlvDn2sjk or SMS your name to 32823	0.431	0.059
· @cocosaofficial IT'S GIVEAWAY TIME! [emoji]RT, SHARE & FOLLOW For Your Chance To #WIN These Beautiful Accessories From @CocosaOfficial [emoji] [emoji] [emoji] #FreebieFriday #GiveawayAlert #Competition #FridayFunFull T&C's	0.428	0.058

Apply [emoji]Competition Ends: 20.07.18 & Winner Will Be Announced Thereafter!		
· Nigeria 1-0 iceland!!! Write it down!#WorldCup #NGAISL	0.428	0.058
· Which saint laurent is your dream bag? Tell us below!Any Saint Laurent order over $2000 gets 25% off and FREE WORLDWIDE SHIPPING!! [emoji]Code is 25ysl! All other Saint Laurent use ìpretty20î#saintlaurent #ysl... https://t.co/mKclvhtKbc	0.428	0.058
· Get building in your home environment with Architectural Lego! Sign up to the Synergy newsletter to win at the @VisionLDN ó stand V119. #VisionLDN #moretimefordesign https://t.co/UZTusDWGR7	0.428	0.058
· @Kim_AgustD @BBCMOTD And thatís salty because of what [emoji] [emoji] [emoji] if you canít see that for yourself your deluded	0.428	0.058
· The Tarot Baby, knows all and seeís all! [emoji] [emoji] [emoji] [emoji]. Will do readings for [emoji] & toys. [emoji] [emoji]. Schedule your appointment now before nap [emoji] time. #TarotBaby #babytarot #babytarotreader #babymagic https://t.co/3IfPBC9QfP	0.427	0.057
NEGATIVE: Dimension 5		
· bITCH MY ACCOUNT JUST GOT TEMPORARILY LOCKED. I JUST HAD A MINI HEART ATTACK	-0.455	0.065
· Free watching of young male gay sex I had him undress all the way https://t.co/ynr3m3YHPc telugu hot gay porn zeb atlas fuck twink videogaysex afganistangayporn gay straight bait bus video full 3gp chloroform muscled gay studs voyeur gay medic	-0.43	0.058
· We love our team development days at Lighting up Learning once every six weeks. Such unified and synchronous thinkers whilst deep respect enables them to challenge each other, including our Director. Even Womble the Therapy Dog gets involved. https://t.co/5KoNLSxUes	-0.427	0.057
· Octavia E. Butler's 'Kindred' remains one of the best books I have ever read. So good I had no idea it was published in 1979 - timeless.	-0.411	0.053
· Update: was meant to be a surprise but just realised my location is still on so she probably already knows Iím here. Fucked it	-0.41	0.053
· I consistently have the dumbest luck, I overslept this morning only to wake up to a message about my meeting being rescheduled to later this afternoon	-0.41	0.053
· and also good morning i had a cool ass dream that i stole a submarine from the us navy and took a bunch of immigration kids and families and a colonel that looked like nate from pacific rim uprising helped me with itand then i searched around to find their families back together	-0.409	0.053

· @RyanHigginsRyan @LarrysComics Target was doing Marvel comics for a hot minute. 3 issues in a plastic sack. 10 bucks. They sat there for a couple weeks, I'm sure. They have now all been replaced with other Marvel merch.	-0.407	0.052
· id appreciate if i was left the fuck alone for the day thanks.	-0.407	0.052
· @DroneOn1 @RealJamesWoods These criminal thugs need to be charged with a crime when they threaten a member of the first family	-0.405	0.052
· @RawStory She seriously took the one good thing sheís ever done with her notoriety, and fucking ruined it yet again with her stupidity. Two steps forward, one tweet back.	-0.402	0.051
· @SarahPalinUSA @northsooman Total liberal Idiot. This cannot be defended. Only reason he is on Fox is for Fair and Balanced. Hell he got fired from a liberal news network years ago. Even they thought he was over the top	-0.397	0.05
· I literally have never been so pissed off and absolutely fed up with something as I am right now in my fucking life	-0.394	0.049
· @williamadler78 @Lin_Manuel The difference here is Stephen Miller is an actual villain. Aaron Burr is only portrayed as one because he killed Alexander Hamilton.	-0.394	0.049
· A bill passed by Rhode Island lawmakers would keep President Trump and any other candidate off the stateís 2020 ballot unless they released five years worth of tax returns.	-0.393	0.049
· The Conners, but where they are all being hunted by a muderous robot from the future.	-0.389	0.048
· I feel like airport bathrooms must be designed by people who have never used a busy bathroom before. "Alright, we definitely need narrow hallways and blind corners. Four sinks and two paper towel... https://t.co/SJjY7CL8n9	-0.387	0.047
· Drilled the weapon barrels and brushed on some black primer, I actually missed so many spots (basically the entire underside was missed when I sprayed them) gonna start on the shirts tomorrow #hobbystreakday207 #hobbystreak https://t.co/mnXoATwPrN	-0.386	0.047
· i just choked on my water for a good 15 seconds	-0.386	0.047
· @wickedlylegit I get that point but Alexa cashed in and she just lost the Title. So she hit back at her. Nia is being an edgy baby face.	-0.383	0.046
· I hate it when someone gives me attitude for asking them to DO THEIR DAMN JOB THAT THEY ARE GETTING PAID $$ TO DO	-0.382	0.046
· I swear I'm a messed up human being inside. No wonder I let certain people into my life. I'm just fucked up.	-0.381	0.046

· @kaiasmommy_888 @principalitysta @edsheeran @StuartCamp They changed their minds about an hour later. But some of us had thrown stuff away by then	-0.38	0.046
· But they already rebooted Queer Eye https://t.co/7eTiZtjFy9	-0.379	0.045
· A little bit about one of our faithful giveback partners: The 24/7 Dad Mentorship is a one-on-one program that offers a relationship with a compassionate and dedicated mentor, education to prepare him for fatherhood, and the opportunity to earn items he might need for his baby.	-0.378	0.045
· Tomorrow, one of my bestfriends/childhood friends is getting married. Someone who i used to make a fool out of ourselves together is now making a life of her own. And i will witness this moment with the rest of the people who made my childhood as crazy as it is.	-0.378	0.045
· My boss gave me his jacket for my birthday I'm SHOOK I'm just a hypebeast now https://t.co/s8TGofHl5p	-0.375	0.044
· Iím just really annoyed that every weekend is so gross out but every Monday through Friday the weather is beautiful	-0.374	0.044
· @StephGrisham45 @FLOTUS For every piece of clothing she buys she could probably feed 100 families. Good point though.	-0.374	0.044
· 4 mins.....a brilliant attacking move from the super eagles rendered incomplete after Iheanacho's back-heel pass was turned to a counter attack....good save from Uzoho to keep the eagle's back line save	-0.373	0.044
· @Malinka1102 @usantidoping He has an allucinated look himself, I think he should be tested...	-0.372	0.044
· Now that I have a tat I automatically win any fist fight with any person with less than 1 tattoo	-0.369	0.043
· It just occurred to me that the only place I'm still considered as part of the youth is in political structures... [emoji]	-0.367	0.042
· Our channel gives total support to the family, we are very hurt by his loss	-0.367	0.042
· Danny Murphy has one hell of a dome on him doesnít he.	-0.367	0.042
· Canít wait one more day to the new ghost adventures I will be glued to the tv	-0.366	0.042
· She just came out of hiding tho... little bounce prolly coming . Sheís usually right.	-0.366	0.042
· I ... got punished last night	-0.365	0.042
· @Lizerenity @kylegriffin1 @foxandfriends @FoxNews He should be fired from being a human being.	-0.365	0.042
· I got lost in the limelight.	-0.364	0.042
· @sugalou_ That is so illegal yk. LIKE THE FUCK IS HE DOING. I ALMOST BUSTED A NUT	-0.364	0.042

· My large veins are all fucked up and bruised cos of how much blood has been drawn from me this past week	-0.363	0.042
· @1MarcMadness I was just bout to say he really started a leagua and for his son to have players to practice against...and they making money off merch	-0.362	0.041
· @Peternoyes4 @PatriciaNPino If more people have more money to spend, as they would under a government spending more on public services and infrastructure etc, then outside investors will be drawn in. Capitalism runs on sales, and sales need customers with disposable income.	-0.362	0.041
· @jerryehudson @VeeVee @morningmika @realDonaldTrump @IvankaTrump I think Vee is a troll/bot here to stir up shit.Many of them are paid per response, so they try to drag it out.	-0.36	0.041
· Bruh Kristen so damn funny , I just can scroll thru twitter and get a laugh from her tweets	-0.359	0.041
· Matthew 25: 40-45 is often ignored by Christians who pick and choose verses as if they are reading a Column A & B menu. https://t.co/c9JDWwliE5	-0.359	0.041
· @___deniseg Empathy ainít gonna keep them out of those camps or being sent back... emotional thinking is counterproductive ... letís do something that can actually make a difference .	-0.359	0.041
· They tried to prevent me from going by confusing me with the time. It did not work. I have tickets. https://t.co/V2eTHZBisV	-0.357	0.04
· Just buttered my toast and went to put the butter away in the bathroom closet. Iím not even stoned right now.	-0.357	0.04
· On Wednesday, my sister, little man, and I took a trip to a local orchard where we picked blueberries. We came home with over four and a half pounds of some of the freshest blueberries... https://t.co/3QCf8GEXwN	-0.356	0.04
· okay so I went to the bike shop and got the part removed in 30 seconds. Then the online place an hour later sends me tracking information	-0.355	0.04
· I'm one of the realest getting overlooked for fraudulent retards https://t.co/DVna49NqfE	-0.354	0.04
· My Lyft driver could get hit, she cute as shit [emoji] [emoji]	-0.354	0.04
· I got a manicure yesterday and my friend convinced me to use gel polish, and now I canít stop touching my nails. These bitches smooth af	-0.353	0.039
· Some douchebag mechanic wanted to charge me $200 for something my friend at his shop just did for free. Definitely pays to know people. [emoji]	-0.353	0.039
· Gatsby hasn't made much noise in the last 15 minutes so either she enjoys Infected Mushroom or the shit she took in her crate made her feel much better.......	-0.353	0.039

· even if you gotta sit in a corner be hated for not talking to them , etc . and sometimes , that makes them want to fuck with you more . dudes . like they did me .	-0.353	0.039
· I had my own infrastructure week..I read a book about Roman infrastructure. https://t.co/GwL5uPJMD8	-0.352	0.039
· Only watched it because my manager bought and gave me the movie so I had no choice, but I give it a solid 7 1/2	-0.351	0.039
· @realDonaldTrump On another note, look into the property appraisal districts- most of us tax payers pay mortgage and when we signed for the loan it included an escrow at that time, yet the appraisal district wants to raise the appraised value every year please, ìstop bleeding usî	-0.349	0.038
· I'm more awake during the day when I barely get any sleep compared to sleeping 6+ hours	-0.348	0.038
· @chrsb01 My current employer was raided in the 90s, detained 300 people.	-0.348	0.038
· @LisaBloor @joansmith1946 I ended representing myself as I was quoted £7.5 thou to read papers and rep me for 1,day in total I was in court 13 days but won	-0.347	0.038
· @boko877 Senpai, I encountered a game crashing bug. When Lusy uses "Tit Down Hold" and Alf suffers from "oxygen deprived" the crashes. It only happens in English.	-0.347	0.038
· @LibsRNutz @busybe_ @teetads @esaagar @NoahCRothman 1. We have not heard from her so we can only assume her intentions and for most of these people it was to ask for Political Asylum2. If she was asking for Asylum she wasnít trying to skip any list...she was trying to be put on a list	-0.347	0.038
· It's a chilly morning but I'm bundled up in the office with my Yu-mi. [emoji] https://t.co/xt9DzKQOJ5	-0.346	0.038
· the angel on my shoulder is telling me to root for nigeria but the devil on my shoulder wants argentina to be eliminated... score draw	-0.346	0.038
· i was told this administration was the best friend to our military. https://t.co/EqUc5a5uVa	-0.346	0.038
· Halfway through my ride, my grab driver reached in his compartment and wore his eye glasses. Like bitch......	-0.346	0.038
· Truth be told, if skills sold Iíd probably be, lyrically Takib Kweli. Trufully I wanna rhyme like common sense. But I made 5 mil, I ainít rhyme like common sense	-0.345	0.038
· Nicki Minaj condemns Trump policy recalling her immigrant experience ëI canít imagine the horror of being in a strange place and having my parents stripped away from meí https://t.co/6Xw2A2ZMeJ	-0.345	0.038
· They fled war-torn Syria and arrived on Bute as refugees. Now, after being welcomed by the community,	-0.345	0.038

they've opened their own business. https://t.co/tweK5i3Vv4		
· @_FCBaller_ Real tears, but only because he's so self absorbed. Scores when the pressure is off and then has a cry.	-0.345	0.038
· I got a sneak preview of the @cawbrighton performance at the Paragon Spectacular. I reckon they may well steal the show @Aurorachief https://t.co/GyprH1WMNz	-0.344	0.037
· @047michelle @MichaelAvenatti ìIíve known Jeff for 15 years. Terrific guy. Heís a lot of fun to be with. It is even said that he likes beautiful women as much as I do, and many of them are on the younger side. No doubt about it, Jeffrey enjoys his social life.î Trump, on convicted child rapist Jeff Epstein.	-0.344	0.037
· i want a one nigga thatíll have these niggas jealous & these hoes mad [emoji] [emoji] [emoji]	-0.343	0.037
· The ultra high note in Say Amen broke me. I love Brendon Urie.	-0.342	0.037
· Frisky pornstar gets her taboo hardcore thoughts fulfilled https://t.co/YNY4fDZ9cd adriana rincon hardcore fake cherie big butt pic black keds fetish filthy blonde arsehole pornstars movies on fbi realitykings moneytalks 3gp www xnxxdreams com	-0.341	0.037
· I was so sleep on carti two months ago then shuffle did me a favor and woke me up	-0.34	0.036
· i have nothing left, no blood, no sweat, no tears... i'm just a dry mushroom don't @ me	-0.34	0.036
· I get distracted by the Internet when Iím on the Internet.	-0.34	0.036
· @bethannieoakley I told my 3 year old about this tweet and he said ìbullshitî.	-0.339	0.036
· I've never in my life heard of a sitting president hiring a foreign intelligence agency to surveil us citizens.	-0.338	0.036
· Always amuses me when somebody called "Bookless" join the library - We also had a Dr Strange joining the library this week.	-0.338	0.036
· U porn casting daybed https://t.co/ms1uChbBmg kanibal hardcore porn pussysexCom wife fucks a dog amateur video www sex hot fucking kaitalin com horny slut gets fucked by two dogs xnxx femjoy free blowjob videos forced marathi fucking videos	-0.338	0.036
· Iëm fucking crying I love bts so much and their new hair colours had me weak,, I wasnët ready for any of this I love them	-0.337	0.036
· @chinolokopana First game 3 times he fucked up on free kicks and instead of letting someone else take it he kept fucking off the clown	-0.337	0.036

· @Dblack_19 @josephsaintval There's a reason he fell that much. Back injuries are no joke. But we'll see. I like Knox a lot though.	-0.336	0.036
· @JoeBrunoWSOC9 @wsoctv Used to be The French Quarter Restaurant at Latta Arcade back the in 90s. They had a great Monti Cristo sandwich.	-0.336	0.036
· @officialmcafee Understandable. Most ppl would`ve been knocked out much longer after having their wiener amputated.	-0.335	0.035
· @MightyBusterBro I too am abhorred by the 3D extremismís and bias on our social media platforms and with MSM. But hereís the silver lining: as the vitriol escalates into lunacy the media platforms serve as a mirror - reflection exposing truth evil hypocrisy. https://t.co/gC7o9llTbf	-0.334	0.035
· For my Leaving Cert in High School we had to write an essay. I did mine on Malcom X. I got an A. And won the class competition. Probably my best academic achievement before college.	-0.333	0.035
· @transscribe @Quinnae_Moon I saw her talk at a Gencon a few years ago I think it was, and was really impressed. Glad I found the twitter to follow.	-0.332	0.035
· @Ifty_ameer20 @cortina190 @CllrBSilvester @YouToryScum1 I worked for the 7 years i was a single mum after my husband abandoned us All my single mum friends including one that was widowed worked too	-0.331	0.035
· @danmericaCNN My least favorite pieces ever. I don't feel sympathy for people who voted to actively harm others, and are now shocked they're being harmed too. We told them, but they voted for a racist anyway.	-0.33	0.034
· @lazchance Reminds me of when the Pistons got Stuckey & Afflalo with their 2 first round picks in 2007. Similar build & game to who they wound up with last night. Not bad considering they were first round picks & going into the night, the Pistons only had a 2nd rounder.	-0.33	0.034
· Mofos act like they scared to fuck with me now... I understand... I'm not at my worst anymore...I'm kinda scared of me too...	-0.329	0.034
· At long last, I've finished my research into the effect alcohol has on physical movement ~ The results were, quite frankly, staggering.	-0.328	0.034
· 99% Ipod the time my phone is never charged [emoji] [emoji] [emoji]	-0.328	0.034

Appendix 4: Extra Tweets Associated with the Twitter Trolling's Dimensions of Linguistic Variation

POSITIVE: Dimension 2	coord	ctr
· Tonight's pin on @11thHour #MSNBC was inspired by a segment on @allinwithchris tonight about Trump playing golf 1 of every 4 days he's been Pres. WH trying to hide this. If the Mueller investigation is allowed to continue, maybe it'll give 45 more time for his favorite activity. https://pbs.twimg.com/media/DSGXnSwV4AE2OUg.jpg	0.683	0.345
· One of David Hogg ‘s personal friends stabs her newborn to death then dumps the body in a neighbor's shed and goes to sleep. #GunReformNow #GunControlNow #GunControlNever #nra #ObamaTookMillionsFromTheNRAToo https://t.co/TAJhlo5viw	0.658	0.32
· These so-called “Journalist” need their credentials pulled immediately, & charged if any crimes were committed. #SorosPuppets #QAnon THE WIKILEAKS LIST: At Least 65 MSM Reporters Were Meeting with and/or Coordinating Offline with Top Hillary Advisors https://t.co/U802GhOqz9	0.655	0.317
· #NewProfilePic Was @willowhalegreen but banned for the truth. 5.500 followers to zero because TWATTER loves Islam and hates patriots.we will never bow to the cult of death & destruction.. https://pbs.twimg.com/media/DRW_G_iW0AAfUJr.jpg	0.607	0.272
· ANDREW BOLT: "Well, this is embarrassing. The crippled Turnbull Government is fighting for its life but has now benched its best two election campaigners: @TonyAbbottMHR and @Barnaby_Joyce." #BringBackAbbott #auspol https://t.co/jqNEdRJx6G	0.603	0.268
· LMAO!! Thanks for the laugh @VanityFair ... Snowflakes melting because #VanityFair did a parody better than @nbcsnl ever could□□□ #WednesdayWisdom https://twitter.com/VanityFair/status/944613785171501056	0.596	0.262
· Can't believe they are not climbing higher to get #Mackiewicz @czapkins - he is the one injured #snowblind & #frostbitten - his blood is on everyone's hands	0.587	0.254

celebrating @ZabRevol return https://www.theguardian.com/world/2018/jan/28/nanga-parbat-climbers-rescue-french-woman … #RIP #sendingboystodoamansjob		
· Judicial Watch filed a FOIA lawsuit against the FBI for records about the removal & reassignment of #PeterStrzok – a former official at the FBI – who was removed from Special Counsel's #RobertMueller's investigative team reportedly due to political bias. https://t.co/MQDTnTlWEW	0.586	0.253
· Smart @BretStephensNYT piece: a large portion of the GOP is super devoted to conspiracy theories: https://t.co/abAS9TBFYF	0.585	0.252
· One week from tonight i will expose @CNN and @NBCNews news for #EpicFail. Happy New Year to all irredeemable deplorables. 2018, cling to God Faith Bibles religion.	0.578	0.246
· The inglorious partisan & devious track record of @AzmiShabana like #sickularists is clear sign that they are a mere front & shield for Jehadi terror groups wedded to breaking up India https://twitter.com/HouseMD1983/status/956808819509665792	0.564	0.234
· Alex Jones has the sads. He feels really hurt by @davidhogg111. Alex will have to double up on his supplements and conspiracy theories. For someone famous for doxing people, Alex is really a delicate flower. SO SENSITIVE! https://pbs.twimg.com/media/DXF6eYRXkAA7pyo.jpg	0.558	0.229
· If @realDonaldTrump calmed North Korea and Syria, inspired Saudi Arabia to tamp down Wahhabism, decreased illegal crossings by 50%, produced back-to back-quarters of 3% growth, created $5 Trillion new stock market wealth, the left would STILL call him incompetent. Oh, wait...	0.551	0.224
· Of all the prominent people who humiliated themselves over pushing a BS Greg Schiano story @SarahHuckabee & @ClayTravis should be the most embarrassed, especially since they both claim to combat "Fake News!" which as my column proves, this story clearly is https://www.mediaite.com/online/the-twitter-mob-is-trying-to-destroy-greg-schiano-and-they-are-100-wrong/amp/ …	0.542	0.217
· Remember, everyone loves wine in their stocking this holiday season. Check out the great @trumpwinery selections this #BlackFiday weekend!https://pbs.twimg.com/media/DPewpgiVAAEeh7Y.jpg	0.539	0.214
· #MLKDay2018 While Liberals & #FakeNews spin @POTUS' words & call him a RACIST, I thank GOD that	0.538	0.214

we FINALLY have a President that's improving the lives of ALL Americans of ALL colors. Obama & his minions were interested in OWNING votes, NOT truly helping people! #FridayFeeling https://pbs.twimg.com/media/DTX7XVVVAAAYalR.jpg		
· Alastair Cook in Tests in 2017: Two double centuries: 487 runs Other 18 innings: 412 runs https://t.co/ZFDm1Kmeq0 #BoomAndBust	0.537	0.213
· #bitcoincore #SegWit supporters remind me of @HillaryClinton supporters more and more every day. We made them obsolete and I suspect the same will be done to BScore. #imwithslow #highfeecoin	0.531	0.208
· The Republican Tax Cut For The Rich Is Careening Toward Collapse In The Senate via @politicususa https://t.co/f24CwSO5I6 #p2 #ctl	0.529	0.206
· At least five companies have now pulled their ads from Sean Hannity's show following his Roy Moore coverage: - @Keurig - @realtordotcom - @23andMe - @ELOQUII - @NaturesBountyhttps://t.co/fsMpagpG5g	0.527	0.205
· GALLOWAY: Theresa May is shelling out millions to combat 'Russian disinformation' instead of tackling terror in Britain @georgegalloway. https://t.co/EaTHgua8ZJ	0.525	0.203
· So RECAP y'all: Jeff Sessions who has been lying his ass off to Senate about Russian contacts wants a Special Counsel to investigate Hillary, Clinton Foundation + Uranium One deal. □ #JeffSessions #UraniumOne #LastWord #Resist	0.522	0.201
· "This team, two or three years down the road, is going to be one of the best teams in the country" - Khalid Hill	0.516	0.196
· Aw Snap! Report: More Major Layoffs Again Coming to ESPN https://t.co/Yew8nUN2bG via @BreitbartNews	0.514	0.195
· The left is obsessed with rape and controlling your girls and boys. The try #WarOnWomen is being perpetrated by the #Dems and #DNC and #Elites. DISGUSTING https://twitter.com/DineshDSouza/status/931395704609693696	0.509	0.191
· Proof that guns can act alone. Poor thing was bullied relentlessly by Democrats and the media. #OneLess https://t.co/cAwCmb2uHQ	0.507	0.19
· Such a privilege and honor to be blocked by one of Grace Mugabe's boyfriends @Hon_Kasukuwere https://t.co/bwJHFYL0AL	0.505	0.188
· He shut down the government over this but now rejects generous #DACA offer from POTUS?Schumer rejects Trump's immigration proposal https://t.co/IeWTW6Yw85	0.501	0.185

· CSK is the smartest team. They are saving money in the auction so that they can use this money to fix the matches later. #IPLAuction	0.499	0.184
· Proud to work for @POTUS who stands strong for Israel and all the Jewish people. #NeverForget #NeverAgain https://m.youtube.com/watch?v=wHqcvxBCm2A …	0.497	0.182
· Church leaders have chained themselves to the gates of Kirribilli House, demanding a group of refugees be evacuated from Manus Island #7News https://t.co/2kGiEVGSTj	0.496	0.181
· gave to @NFL 12 Stop Kneeling 11 Trump insulting 10 Fans Boycotting 9. Ratings tumbling 8 Goodell mumbling 7 Tix prices dropping 6 Jersey burning 5 Sponsors running 4 Owners crying 3 Stadiums emptying 2 Empire crumbling 1 America WINNING	0.495	0.181
· Donald Trump walking with Putin looks like a giant toddler walking with his daddy. #saturdaymorning #APEC2017 https://t.co/QEpJPjLcbh	0.487	0.175
· @FoxNews @TomiLahren @POTUS @HillaryClinton HRC is the most corrupt woman on the face of this planet. She made off with 145 million selling Russia Uranium. She made of with hundreds of millions from Saudi Arabia. And another couple hundred million from her pay to play schemes as sec of state. She needs to be locked up	0.482	0.171
· #MichaelBennett feels like a slave making $11 million per year while men who serve our country don't make even 1/4 of that. Just one more reason to #BoycottNFL ! https://t.co/ziBuEwuFnH	0.48	0.17
· Liberals want to #CancelVanityFair because the publication had the COURAGE to challenge arch swamp-demon, Hillary Rotten Clinton. Vanity Fair is only reflecting reality... https://pbs.twimg.com/media/DSDUAcPX0AAJ69B.jpg	0.477	0.168
· JJ Watt raised $37 million for Hurricane Harvey victims. 37 MILLION! But Kaepernick refused to stand for our national anthem (a year ago) and is Citizen of the Year. Right...	0.476	0.167
· One pic worth a thousand words. Turkeys love of pedophile jihadists makes sense now... Birds of a feather #Erdogan #Afrin #AfrinOperasyonu https://twitter.com/fate_tahir/status/968647638039089152	0.475	0.166
· Not if they use Sen Al Franken against the DFL in 2018. Wellstone took the US Senate seat when incumbent Rudy Boschwitz backed sex harasser Grunseth. #mnleg https://t.co/lRxUu327ym	0.474	0.166
· @brithume @Nvehecnycrrcom1 @AndrewCMcCarthy One of the greatest threats we face right now is Rep. Adam Schiff interference with the truth in	0.472	0.165

our elections and the general eroding of our democracy due to (some) Democratic complicity with Rep. Adam Schiff's dangerous rhetoric. #TickTock #fisaabusememo #FISAGate		
· RT @SkyNewsBreak Formula One race driver Lewis Hamilton found dead at London home https://pbs.twimg.com/media/DSFFnliWAAA1Fws.jpg	0.471	0.164
· Indiana Nurse Under Investigation for Tweets About Killing White Babies #PROTRUMPMVMT https://t.co/AuzQD4J74A	0.466	0.16
· Would Have Been Easy 3-0 Loss Had Rahane & Bhuvi Been Dropped Yet Again In 3rd Test. Would've Been A Series Win If They Were Selected Together In The 1st & 2nd Test.Hope Virat Kohli & Team Management Learn From Their Bizarre Mistakes. Congrats Team India□□□□□ #INDvSA #SAvIND	0.465	0.159
· @DVATW Thank heavens the #Scottish electorate are waking up to the fact that the #poisonous #Nicola #Sturgeon has had her day. @NicolaSturgeon #lost her #indyref & lost seats to @scottishlabour & @ScotTories in 2017	0.465	0.16
· @realDonaldTrump 2017 was the best Economic performance for America is DECADES, thanks to President Trump! 2018 is going to be even better! #MAGA	0.461	0.157
· YES‼□ REAL GAME CHANGER! Liberal Followrs OF DNC Could NEVER Wrap Their Heads Around This! NOT SUPPORTING TERRORIST IS THE RIGHT THING 2DO! This Crooked CIA Program Jst Came 2 A CRASHING END, TRUMP Had It CANCELLED!□□ @realDonaldTrump http://youtu.be/N-EUo41iLb8 via @YouTube https://t.co/mXeTQnIUsT	0.458	0.155
· Absolutely not. Source said Currie had his guy signed and there would be an announcement today. That was correctly called 16 hours before national media. Then the glorious uproar. https://t.co/QW2sdoaIvf	0.457	0.154
· This probably should have been established before he was allowed to twice take the stage, once in a rally speech before the townhall, and blame millions of innocent families for his incompetency. https://twitter.com/CNN/status/968454239516282881	0.453	0.151
· In the first half the Knicks made LeBron eat his words. Knicks by 13. LeBron, 5 turnovers. Ntilikina, 5 steals. But long way to go.	0.453	0.151
· This administration has all the wrong priorities. They're trying to increase support for corporations and the richest 1%, but slash support to empower women farmers and feed some of the world's most vulnerable populations. https://t.co/6M6b9A5MtW	0.453	0.151

NEGATIVE: Dimension 2

· @Salon I'm white. I don't see it. Morons. How is he a racist? I don't think you know what racism is.	-0.527	0.205
· @RomeDoesIt i don't know why you're still so concerned about a troll tweet. i hate idiots like you.	-0.493	0.179
· @NEERAJ_AGARWAL_ @barmanamar1976 @CFBKEW @DeanKo @WeAreWakinUp @phiroc @_Gravity_Man @Theflateartherz @VerumBellator1 @FlatEarthCity @ItsFlatFolks @facebones777 @HomeoReikiDogs @SpeakToMeInDots @Its_Stationary @GodofGreen2 @nutsyLFC @Spacehehehe @ADalassio @VickyAlam18 @Th3NewMoon @BadBuc99 @ericdubay @IronRealmMedia @catomilla @IllCity_Luck @jeranism @TheWrongQuest @jaredvc @mode23 @hugh_bothwell Just because you don't understand something doesn't mean it isn't true lol.	-0.479	0.169
· @idkasuri @Fabulous_IK @ummesalaar @Paracha_Pk @kambohguyforpti @StaunchInsafian @FauziaKasuri @BBhuttoZardari Oh we know what you're talking about. Get it off your chest anyway if you want.	-0.472	0.165
· @helenlooise Did you have some point to make or whatever? I don't know what you're on about.	-0.462	0.158
· @Johnlee333333 @arthutchinson2 @realDonaldTrump To someone like you who doesn't know what truth is, I'm sure it seems that way.	-0.461	0.157
· @ElenaTayTay @tariqnasheed Why don't you tell me how oh person who can't even spell? I'm sure your ideas are brilliant. □	-0.46	0.156
· @ryanthillparody @BarstoolBigCat Also, you don't need to be dick. I'm pulling for him as much as any Badger fan, but that doesn't mean we can't criticize him for being bad.	-0.458	0.155
· @canadopia @pewdiepie @nytimes Honestly offended. So just because I'm not 'A' that must mean I'm 'B'? I thought you were more than that.	-0.454	0.152
· @TheOneTrueMin @Dilbert22022194 @BeyondPhere I'm assuming you're British?	-0.452	0.151
· @DaRaiderz4Life @rlfcars @FoxNews How can you say I don't know anything about the strategy? It sounds like you're projecting…	-0.45	0.15
· @LexieMatheson @richardhills777 @barrysoper @paulabennettmp Really? Well then I'm a helicopter. So do you acknowledge me as a helicopter now?	-0.446	0.147
· @BarbaraLandree1 @malenroh1 @MattJ87220412 @FoxNews @POTUS Nah, I also don't claim to "know" things like you do.	-0.439	0.142
· @AlwaysThinkHow @huckfinn22 @peterdaou No, you are divisive and that's exactly what you're trying to do. All of your rhetorical nonsense is just lies. Do you even know why you don't like him?	-0.438	0.141

· @ang_zone_vile No you did not call me on it. You did not call me on what I said you called me on what you thought about what I said without even acknowledging that your words are not what mine were.	-0.423	0.132
· @TaylorC07519097 @NtmAjBushey @Kristen_Taketa @twitjb Ok you must not understand that ignorance is an epidemic.... if my replies are so irrelevant why do you keep responding sweetheart	-0.423	0.132
· @KingAcer33 @tariqnasheed @benshapiro I knew you didn't have evidence and I'll consider this a win. It's been real.	-0.421	0.131
· @BIGKIDXAVIERLOL @CONNORisSWEET @baxbooksdeux @Ruadhain_K wow I can't believe you're being biphobic like this to me	-0.421	0.131
· @mohawkmattzyr That's literally what I said. I quoted it word for word. Lmfao you really are desperate	-0.414	0.127
· @My57ChevyBelAir @SpencerChretien Ok but that's not what "white privilege" means so your pride is misplaced.	-0.411	0.124
· @Horror_Mistress @ZaackHunt Um, first I said too so your tweet is invalid. Second, it doesn't help that it looked like you were talking to yourself.	-0.407	0.122
· @TWilder86 We get it, you are easily entertained but its not as easy for to think consciously.	-0.405	0.121
· @ImpeachOranges @Dave_Mitchell73 @Cernovich @realDonaldTrump Don't expect anyone to give you answers when it's clear you are too lazy to learn the truths on your own. smh	-0.401	0.119
· @Rachie_Rach__ @realDonaldTrump @RealSlimSupreme @GAPeachMEG @MistaBRONCO @hgjoz @dt_TruckinOn @RottenToesJones @SpayMsm @LT51552424 @_Suga_Glida @jensuz73 @joej2020usa Bitch we know you aren't real lol	-0.398	0.117
· @Hon3y_Be @vlynxy1 You definitely have no idea what kind of person I am. To say otherwise is to admit your ignorance. Pointing out your hypocrisy isn't childish. Being a hypocrite is childish, so you should really try to stop doing that. It would help if you would grow up.	-0.397	0.116
· @Illuminati4all @cinderella2b @DavKat43 @The_Trump_Train @realDonaldTrump I assume you've seen their Tax Returns or W2 form? Don't be a racist. It's ugly.	-0.389	0.112
· @sjimmyp @garygil58645417 @GlobalBC Why yes, yes I do...but you don't.	-0.386	0.11
· @RedPandasDaily i dont think youre getting my point, you can have any views you like and its totally okay, but dont be vocal about it, like tweeting about "blocking nazis" and such its unprofessional after all, and im sure we can both agree that the red pandas are more important than any of this	-0.386	0.11

· @Maurice22015378 @miss_speech @realDonaldTrump Natural English? Lol... Because English is so original... And you're about as hilarious as cancer... I find it amusing oh how funny you think you are	-0.386	0.11
· @Seeds81Planting @RRRDontTreadOn @USATrump45 @unconcious0 @cjdtwit @luluHru @knkcattle @1ofthegoodguyz @trbrad62 @Alice00581238 @uniquedeehan1 @TPrincipata @tellilikeitis @pjbjr2485 @RoryGilligan1 @momof24u @LovesTrump45 @ladydiblu1 @BeerMeMarge @PIRATEDANTRAIN @SanDiegoRuthie @LynwoodTalks @OdinMo @ShoreyMichael @RUSTIMCCOLLUM @NinaGrigsby @CAoutcast @SKSSKanz @TechQn @Courtneykh24 @CudaDebbie @ruby58293 @wanttruth @StrongShepherd_ @45isMyGuy @bbusa617 @GeorgiaDirtRoad @ArizonaKayte @jimlibertarian @DeplorablAnnJoy @TrumpsBlonde @HippoCovfefe @spacegirl1 @KoalaFan3 @izonorion71 @PVHenryConLLC @Blondi1210 @BlueColossus ad hominem attack + you're * your ...and I'm not original???	-0.384	0.109
· @CheriJacobus @SeanHannity__ @PMorici1 Excuse me Cheri, are you for real!? Another Dumbocrat who doesn't know what they're talking about!	-0.382	0.108
· @AllPowerFades @scorpixon Sorry, that's fair. So which of that list of philosophers - all of whom I've wasted many hours reading - is the support column for using "that's just the way it is" in lieu of actual proof? Before you answer tho: I will not accept any attempts at actual proof, so don't try.	-0.378	0.105
· @vilevillainess @dediane1956 @veterans_i @brat2381 @Navyvet270 @leigh_proper @imahaider @RileyChildrens I'm white and not particularly accomplished. I think both of y'all are racist.	-0.377	0.105
· @DestinyandBruce @EvOConnor15 @TimRunsHisMouth @DonaldJTrumpJr @POTUS @realDonaldTrump @FLOTUS Oh dear lord you all need help. A) If he didn't want to give up what he had he wouldn't have fucking ran for president B) Stfu if u don't know what you're talking about	-0.374	0.103
· @VictoriaBanvil2 @DavidBegnaud What are you a fascist ? You decide if someone can reply or not? Don't confuse me with a colonial subject... that you tell what to do or not.	-0.374	0.103
· @JerrryGrey @AnnastaciaMP @DrAnthonyLynham @jackietrad @QLDLabor @LiberalAus @The_Nationals Oh stop crying, we know you mutts like to play the victim card all the time. Seeing as you are blind, deaf, as well as dumb, have a look back to when I first said "Typical mutt communist reply". There you go, I've done the hard work for you, seeing as it was too tuff for you.	-0.373	0.103

· @Auroraknite @EmilyLindin @jasminegjackson @jeremypiven @AsiaArgento Are you saying I'm falsely accusing Emily of something? I'm not, but if I were, it's a price I'm willing to pay.	-0.373	0.103
· @JMTheAtheist @MMattstofferson @Breaking911 You're pretty stupid aren't you?	-0.373	0.102
· @TheReal_Gabi So what you're saying is you're dumb?	-0.372	0.102
· @PatriciaAnn3225 @cloverjag Oh, so you'e high on opiates.. that explains your erratic behavior. I'm still not sure about clover though.	-0.371	0.102
· @sbhopper8 Lol It's not about knitting at all. But don't deny me my fun.	-0.37	0.101
· @IFThunder Awe was that supposed to hurt me feelings?? Not happening! □ I don't have delicate feelings and i'm not emo But I do Block idiots like you!	-0.369	0.1
· @Cat_MarqueeLV @outandaboutjc1 @DB701 @cmr4to @Rorenado @ann827 @AP I was trying to talk politics like an adult. I guess that's not an option. You know pointing out where we agree and where we misunderstood understood each other. But I guess you don't want to acknowledge that we agree on a lot of points because that doesn't fit your narrative.	-0.367	0.099
· @ericgarland Do you think it's wise to publicly accuse journalists of treason and threaten them with imprisonment? That seems insanely libellous	-0.367	0.1
· @lindahillmorris @bex0760 @cucch327 @baldy1004 @RepAdamSchiff @EBlumberg11 Yeah, actually it wasn't a good point but hey you are a lib so I know logic isn't a strong point.	-0.366	0.099
· @mohawkmattzyr You won and you have no idea what game we're playing. □ If you want me to take you seriously you have to actually prove statistically that police are bias. Gl with that	-0.365	0.098
· @SamTSnelling You're not exactly being sensible yourself... so...	-0.364	0.098
· @flm22 @casara66 @fawfulfan @DebraMessing @nycbubbles @LeahR77 Ummm faithm, there's no such thing as a full term abortion.At full term, what you have is a baby that can survive on its own,outside the woman's body. No one has an abortion at full term, don't be daft.Argue w/facts if u happen to have any. I remain skeptical but will await word.	-0.363	0.097
· @dinahkiwi @ShadowhuntersTV it wasn't an opinion when you said "us all" unless us all means just you?	-0.363	0.097
· @ZiplockGaming @PeterSweden7 Well done. A speed reader? You have learned nothing. Don't bother communicating again until you are educated.	-0.363	0.097

POSITVE: Dimension 3	coord	ctr
· @VictoriaBanvil2 @DavidBegnaud What are you a fascist ? You decide if someone can reply or not? Don't confuse me with a colonial subject... that you tell what to do or not.	0.588	0.332
· Since Thomas Wictor has me blocked (don't know why) I am tweeting this valuable information he has provided. Handwriting from year book Roy Moore accuser provided, does not match & person got Old Hickory House wrong. Please retweet Thomas or me. I want this information exposed!	0.582	0.325
· @billings_steve @VigorousRaDiCaL @lorenzabraham12 @GhosTNinjaFtW then do me a favor don't say i said something about you when I didn't. If I wanted to say something to you @billings_steve I would have brought you up	0.582	0.325
· @AnnCoulter @SUPgrlCaroline @realDonaldTrump Trump, with all the idiotic choices on immigration has done what the whole world could not do, he has turned me against him. Let Mueller lock him up for lying to us on the wall, and amnesty! At this point I really don't care if he gets impeached or locked up!	0.571	0.313
· @ESPNFC When did USA become a team to miss @the world cup? Don't Make me laugh! They can miss 10 World cups for all I care! Won't be missed!	0.547	0.287
· @amandablount2 Liberals are constantly trying to tell me what to do, and forcing me to do it via government force if they can. You don't like it very much, do you?	0.545	0.284
· @BasedTXPatriot @EQFoundation @Securityconcern @networkradious @TheHoneybee_ @StarSpangled9 I understand why people sell drugs, assault, Rob, murder in some cases.. But pedos?? No gain at all, only hurt. If we put them in prison with no PC they will be dealt with incredibly swiftly.. But the prisons protect them so weirdos like this guy don't cry and complain	0.499	0.239
· @womanINtransit @johncusack 2) Drastic measures as in conspiring to topple him? Says a lot when we have to overturn the democratic will of the people with an establishment coup because they don't like what he's doing. What happened to accepting the election result and the rule of law?	0.498	0.237
· @IngrahamAngle Laura! Don't you think these kids were organized by someone? I believe they were given these talking points and fired up by certain people that want to push this narrative	0.495	0.235
· @TheNettieRhodes He did bring us forth and is empowering us!!!, We are called "Deplorables" We ...	0.49	0.23

can't... stop... WINNING! You silly goofball. MAGA-KAGA. Save this tweet sweetie. I will ❤		
· This one surprised me. I like Joe Rogan... but 1.6 million fake followers? He didn't need to cheat. WTF. @joerogan @eddiebravo @BrendanSchaub @ChaelSonnen @MMARoasted @TimKennedyMMA @dc_mma @jeffwagenheim @MMAjunkieGeorge @MMAjunkieJohn @MMAjunkieSteven @MMAjunkieMatt @Benaskren	0.486	0.226
· @gingerfossum @bigdinkel @realDonaldTrump He most certainly did. You realize Donny's been married 3 times right? I don't need a news outlet to give me a reason for that.	0.485	0.225
· @carl5480 @brithume @Patterico My point is, why now. This guy ran a tough primary and you're telling me it didn't come out then? Only when a left leaning newspaper reports it? What if he didn't do it? Then who's being persecuted now?	0.485	0.225
· @glenpen60 @jukesgrrl @xorbanana @Michael46291030 @PurlLeslie @RepAdamSchiff Why do liberals feel they have to lie to get people to listen to them? I know liberals tend to believe whatever they are told, but I wonder if your cause us so good,why do you need to lie about it.	0.481	0.221
· I dunno. Why don't you ask some "black people"? Just walk right up to them and repeat the tweet. Oh, and pretty please, use the finger air "scare quotes" for special emphasis. And pics or it didn't happen. https://twitter.com/Mel_Ankoly/status/930124076969992193 …	0.478	0.219
· @LindaMasonJar @WVGovernor Why do people need private insurance? We all get moved onto a single payer system if we live to 65. Ask an old person if they want to dump their Medicare and go find private insurance. Stand out of reach, though, you might get slapped.	0.474	0.215
· @ResistResistR U don't know how good U have it.Obama nearly destroyed us. Pres Trumps given us respected international relations,Vet Reform, more jobs, bus.regulations lifted, economy growth, bordercontrol,ISIS destruction. Impeach?Cuz you don't like hs personality?	0.473	0.214
· Listen Up Libtards: "DON'T F-ING PC POLICE ME" https://t.co/S4aL8Jc7SE	0.465	0.207
· @DarkAngel_USA @FoxNews Who said anything about parties there superhero? Ray Moore? I don't support him. Your life sucks that bad that every day you wake up and place people screw ups on a political party? Dun duh duhhhhh	0.458	0.201
· @IRdotnet @realDonaldTrump @jamiejmcintyre @dcexaminer Remember when we had a President who	0.434	0.18

didn't need to brag about accomplishments? Yeah, I remember Obama too!		
· Whoever is shipping #Reylo I am questioning your judgement. No redeeming the guy, he doesn't want to be saved and Rey shouldn't have the responsibility of his dark soul https://twitter.com/MyFandemonium/status/946098729341165568	0.431	0.178
· First though, do you have a woman in your home? A sister or a mother or a friend? Be a real male. Give them a hug (ask consent first) and say "I support you. I support #repealthe8th". Gestures matter to our woman sisters.	0.428	0.175
· @Scavino45 @POTUS @realDonaldTrump Boo we don't want him back. Keep him there!	0.426	0.174
· @freespirited_p @DawnButlerBrent @emmadentcoad So @EmmaDentCoad has blocked me, perhaps she doesn't like black Tories reading her tweets. I wonder why. https://pbs.twimg.com/media/DPmDsZMW0AE7yIt.jpg	0.42	0.169
· @WJ_Armstrong @canokar @06JAnk Well then you should never complain if you ever get silenced or censored for simply speaking out, w/o any harassment, insults, etc. What you waiting for? Block me already!!	0.416	0.166
· @MaryMister7 @ltucker8044 You want me to swear at and threaten you like you did to me?	0.414	0.164
· @FactsMatterHere @EdanClay @Melinda15858273 Sorry the photo of him and Ms. Huffington tells me all I want to know. If an ignorant person saw that he might think Franken condones that behavior and justifies harassing women.	0.413	0.163
· @DasitBoo @CBSNews Let me know how much of your money you want me to take	0.41	0.161
· @BradfordNims @PaulsEgo Why can't you and all the pro gunners just acknowledge the reality of your position? You don't care about people being gunned down in churches, schools, and other public places. You think your personal "Freedom" to have an AR15 trumps the rights of the others to live. Be a man.	0.409	0.16
· @grandedamegria @foxandfriends @foxandfriends @trumps_feed I just love America if you don't like someone's opinion just call them a bot	0.409	0.16
· @Watcher4321 @rahulnag @HomeLoansByHDFC If you really think banks are cheating, why take a loan from them damn it. You need them when you need money and blame them for your laziness to refinance your loans. GROW UP!!	0.407	0.158
· @NRAVikki @Scherazad100 @sue91sue @GCando1 @CaroleW008 @jordansdiamonds @MikeMason830 @HillaryClinton Carry on preaching to	0.407	0.159

your choir. I don't hate immmigrants.I don't hate brown or black people.I love children. I dont think poor people should suffer or be stigmatized.Refugees deserve our help. Corporations don't deserve tax breaks.Disabled and lgbt people deserve equal rights.		
· @CheriJacobus @SeanHannity__ @PMorici1 Excuse me Cheri, are you for real!? Another Dumbocrat who doesn't know what they're talking about!	0.407	0.159
· @berkfran @maggieNYT @rubycramer Back then, her people wanted to can him. She said “let him@stay”. https://t.co/wfIiColkzr	0.405	0.157
· @GenDesignInc @CockOfTheWalkDP @BetoORourke @tedcruz Go ahead and waste your time. I don't care. Texas will re-elect him. Then you can cry and write about it.	0.404	0.156
· @that_nocoiner @indystar We have no moral high ground here and we missed an opportunity to have an important and necessary conversation. Instead, folks like you can feel vindicated for not liking "those people" because, well, they don't like us either. https://t.co/YK235qMAui	0.401	0.154
· @FLOTUS @NIH @TheChildrensInn Did U have your red bikini on underneath your coat & a 'concealed carry' weapon ready to defend the kids if anyone who had legally bought an assault rifle & loads of bullets came in to shoot the place up?□ I prefer the demure & very ladylike @MichelleObama a GENUINE 1st Lady. https://pbs.twimg.com/media/DWuMVcUWkAE6JRI.jpg	0.395	0.15
· UConn just lost by nearly 40. What positive statement would you like me to tweet about them? https://t.co/if3yGLkdmo	0.392	0.147
· @ABOLISHWELFARE @frenchjonathan6 @bbusa617 I don't want to be followed back. I stand by my beliefs not by what some preacher taught me. I have a brain of my own.	0.392	0.147
· I don't dislike DC or Zack Snyder. I'm just still waiting for him to make a good movie. https://t.co/3dEOPd2gif	0.391	0.146
· WARNING: Do not show this tweet to your kids, because they may lose brain cells reading it https://t.co/1Q1Sqrrpv3	0.39	0.145
· @giddieupbitches @Early__May @TaylorEdwards99 @AJDelgado13 The point Your head Do try to keep up. Have you read the book? How can someone be brainwashed for calling attention to the fact AJ hasn't read it?	0.389	0.145
· @justgord Sure, big blocks work, until they don't. What are you going to do when we fill up all 8 MB consistently? Continue to increase the block size until we	0.382	0.14

start having all nodes download terabyte size files? #BCHScam		
· .@NYCMayor wants to address #climatechange – but not if it means giving up his SUVs and helicopter rides. #ExxonKnew http://bit.ly/2ATdAvD	0.379	0.137
· @davidhogg111 How much of THIS KIND of BULLYING did you do to Nickolas Cruz? Did you ever befriend or have lunch with ANY "outcasts" in your school? If not, then perhaps YOU should look to yourself for "solutions" --- Maybe YOU could have saved your classmates, by simple acts of kindness.	0.379	0.138
· @CBJOHNSON143 @realDonaldTrump Again, you still didn't tell me, plainly, where did Trump break the law?	0.379	0.138
· @jeffrey_hatwig I don't understand how you being offended you were proven wrong makes my comment unrelated when it is 100% related and calls out your bullshit?? Btw, what the hell does @$$ mean? I don't speak third grader.	0.379	0.137
· Now we get to ask a new question: Why does Jared Kushner still have a downgraded security clearance? Also, why hasn't #Kushner responded to this simple question from Congress Members: Did he have any talks with #Saudi foreign nationals about the troubled loan for 666 Fifth Ave? https://twitter.com/kylegriffin1/status/968589077724516353	0.376	0.135
· @plainviewsue @zebstwit @JohnPaul_USA @realDonaldTrump @PressSec @SarahHuckabee @HillaryClinton @NFL you dont even know him plainview f sue my fam worked for him everyone of us minority dont u care about israel #NOT	0.376	0.135
· @BlackHaven3 @TMZ Thanks, but I didn't need your permission. Your tweets had a point? Hmm. Must've been lost amidst all that poor punctuation.	0.376	0.135
NEGATIVE: Dimension 3		
· @BobbyMartin044 @TBama23 @NFLResearch Its the whitest team in America. since 9/11 the Patriots have been Unstoppable. New ENGLAND, PATRIOT, red-white-blue, the HAMPTONS, POLITICS, it's the whitest area of the country, the coach is white, QB is white, RB is white, WRs r even white! There is a white man on the helmet!!	-0.547	0.286
· @AlterSol @brianherman @pma19722 @RightlyNews @Fuctupmind Facts matter: Trump has best stock market performance in 20+ years. Lowest unemployment in 17 years, um, oh yea, that's Better than Obama ... EVER!	-0.542	0.281
· @thebuddhacat1 @IanMCohen @azmachman @John_AKA_Becker @SassBaller @CNN	-0.49	0.23

@BarackObama @POTUS Stupid at its best coming from a "nasty Buddha cat".		
· @THR Well... The Room is pretty horrible... But at least it's better that #LastJedi https://t.co/r01rwV7o8f	-0.483	0.224
· @breton_anne @DestinyandBruce @SergeFauchet @EvOConnor15 @TimRunsHisMouth @DonaldJTrumpJr @POTUS @realDonaldTrump @FLOTUS Nope. Obama wears the biggest liar title well. But that's okay. Sit back and enjoy the ride.	-0.475	0.216
· @dcombs065 @Charmie910 @bonnieotterson @DineshDSouza @realDonaldTrump Jones thinks that stabbing an infant in the base of its' skull, as it is 90% delivered, then sucking its' little brain out is ok. Partial Birth Abortion is far worse than any of the false accusations they are bringing against Judge Roy Moore. Jones sucks but Alabama knows better!	-0.465	0.207
· @WhoWolfe That's funny. Trump is also a loud mouthed obnoxious fat pig.	-0.453	0.196
· ANDREW BOLT: "Well, this is embarrassing. The crippled Turnbull Government is fighting for its life but has now benched its best two election campaigners: @TonyAbbottMHR and @Barnaby_Joyce." #BringBackAbbott #auspol https://t.co/jqNEdRJx6G	-0.45	0.194
· @threejuniormnts @RealtyVirginia @RyCliffordCares @JackPosobiec That was like bringing a gun to a knife fight. No pun intended. On the bright side, you've already written the galley for "Boycotts for Dummies," so not all is lost.	-0.445	0.189
· @RattoNBCS That's not a serious comment, right? The best minor league game isn't even close to as good as the worst Pro game.	-0.436	0.182
· @carlos_valencia @wmmII88 @realDonaldTrump You obviously are biased and are intentionally ignoring all the facts exonerating Trump. Mueller, rosenstein, comey, Lynch, Clinton, DNC, McCabe, deep state all clearly colluded... this is the most obvious coup that has happened in the USA. Sad to see the media ignoring facts.	-0.432	0.179
· @Vet4MAGA Fake news snowflake. Delta is stronger than ever.	-0.431	0.178
· @BirdyLovesIt Sex with strangers, or camming: it's all pointless. Porn is free, and dating or one-night stands are better than any escort. In a few years camming sites will go extinct.	-0.429	0.176
· @jmcg134 @JJWinter62 @KarlDodsworth1 @nickpertom @SkySportsNews In summary, How De Gea is crap. And your man that replied is a sexiest pig. That's all.	-0.417	0.167

· @EdBattes Those were some ugly ass women. Ember Moon finisher is stupid. But overall pretty good show so far	-0.414	0.164
· @acton_rosacea @NBCNews And sex predators are bad, im sure you agree it was horrid hillary defended a child rapist in court saying the child fantasized about older men even married a predator.	-0.413	0.163
· @iMau5xx @JohnnycarlsonxD @DepressedNets @lunaeverything @HasimovicEdo @FortniteGame No lol one of the best games out right now it's very well made and the concept is just so interesting.	-0.411	0.162
· @wtfkenneth @ryanxcharles Hopefully the devs were polite about it though because people skills are more important than technical skill in your brainwashed community of lemmings	-0.407	0.159
· @Leamons71 @ChristyCallah15 @IngrahamAngle @trumps_feed @realDonaldTrump Not everyone, just the nimrods that try to push a Russian coup as okay, fine, nominal. There's no way in hell a normal American citizen thinks this is okay. Ergo you're not a normal American citizen.	-0.403	0.155
· The @realDonaldTrump hotel in DC is far emptier than its competitors, and charges far more. @cnn w/ great new data. https://t.co/7h8FRQgpeb	-0.402	0.155
· @realDonaldTrump The #MAGA movement is bigger than any other in American History!	-0.401	0.154
· @HeySaum @RealKidPoker There is no data that supports the premise that women are more emotional.	-0.399	0.153
· @ArmiinWoozworld Ok goodnight. Pressed delusional navy. Beyonc√© will forever be better than Rih. Bye.	-0.388	0.144
· @superkruger @OrwellNGoode the SA constitution is a pile of shit and you're stupid for thinking it was ever good for anything.	-0.384	0.141
· @heyyjass @Defiant_MIKEL Bugging phone is perfect, Apple >>>>>> everything else	-0.383	0.141
· Tom Brady continues to play at a higher level under more difficult circumstances than Carson Wentz. But there's no way Brady wins MVP. Too much jealousy and resentment. He's too great for his own good. He's already won too much even though he deserves to win again.	-0.379	0.138
· @DatDude126 @LisaGau92353626 @WWE @JohnCena @AJStylesOrg Uh, literally the most unimpressive character in dragonball super? Ok, I'll be sure to remember. Seriously, yamcha was better	-0.376	0.136
· @djeffery547 @NoIdiocracy @HouseOWar @M_Turner65 @rosebuddlove @icklepepper @Salsa_Sommelier @kellydraper @altDoD_ @DLoesch @benshapiro hahahahahahahahahahahahahahahahahahahahahaha	-0.376	0.135

hahahahahahahahahahahahahahahahahahahahahahahaha hahahahahahahahahahahahahahahahahahahahahahahahaha hahahahahahahahahahahahahahahahahahahahahaha....... your English is way worse than the undocumented folks I know.		
· @XMalice @ResistanceRosie @matthewjdowd Sex is optional. And none of you are very good at it anyhow.	-0.375	0.134
· @RealJamesWoods I am a grief counselor. This is the farthest emotion from grief that I have ever seen.	-0.374	0.134
· @OhhDee92 Harden never made the all star with OKC dummy. □ no knowledge.. this is too easy □□	-0.373	0.133
· @Jersey_Gulls @Goss30Goss @SheRa_Princess_ @DocPeteyJ @lockedmith @veterans_i @OleVetUSAF it's a blatantly obvious fact and everybody knows about it. any Flyboy, Jarhead or Squid will be able to tell you the jokes	-0.371	0.132
· @GerardWhateley "It's an old-fashioned coup, but it's absolutely deceitful, duplicitous and distasteful" □	-0.368	0.129
· @CatsYouCanine Luka Doncic, DeAndre Ayton better and so is Mo Bamba. MPJ would've been. Only in college Ayton so much better.	-0.366	0.129
· @SilverAttack @CitizensFedUp @shuboogie @BroMark13 @10MinutesaDay4U @Eviljohna @brocanteuse @PWM62 @BklynDin @SailingCarol @RFreeba @NannieWulforst @CrippledJerk @WendyCarey2 @ooshdesign @SpryGuy @MarenHofstad @Deemoney521 @biegenci @WalktheTalkBern @AsperGirl @JonJusth @DeniseO6229 @GOVTWINE @TruthEqualsFact @bannerite @doctoralexa @MoniqueMellon @CaptainsLog2O18 @alliewen21 @_WeStandUnited @grandoftwo @jglewis1983 @bmr0954 @kenner3616 @charpig @SimpleWonders82 @Gabbiedrice78K @PurpleReign14 @anonycraig @tweetMalena @ChicagoMGD_SD @chance4gardener @SchMieke @Loynaz15 @TheGodlessMama @fernaleaha056 @morganarae @62WalterP @PamelaGlasner Hillary's approval rating os lower than Trump's. Fortunately, you Hillbot assholes are such a tiny minority that you can't divide the party. So go fuck yourself.	-0.364	0.127
· Wikileaks is real news, which the fake news media HATES. It's easier to sell counterfeits when there are no genuine articles around.	-0.363	0.126
· @JoezMcfLy You have a point. Definitely over-achieved. If that's possible with a 200 million dollar payroll.	-0.36	0.124
· @YungFraudGod @GQ__Jackson @ImReallyTrilly @no_cut_card @zoeveli Actually you're wrong about the second part. I teach men to fish. So basically there is	-0.359	0.124

nobody more virtuous than me. I'm there at the top with Jesus.		
· @Gardeniagal4 @EmmaGPaley @miche371 @mmelgar09 @OSUCornboy @Plasticdoe @LilEarthling369 @jim_herd @Sheeple101 @tyoung_5 @JUVerastegui @viva_lala @doritmi @LTock @marcdraco63 @LauriLinnea @DanaElizabeth69 @Orangesec333 @PharmaNemesis @eTweeetz @janem1276 @LaLaRueFrench75 @CSavamom @Just4TheCause @qtbeauty @FarmgalMom @TheFrankmanMN @kidoctr @jkellyca @agargmd @ghoppe @Organic_Mumzy @MilanovNina @KristenJayne1 @regina1775 @and_kell @AlokPatelMD @SaveTWRadio @Charbrevolution @Vbalance03 @A_Silent_Child @Cloudhunter @badzoot7 @StopVaxxedLies @kenjaques @steffieschiltz @nicolasDenver @science_guy5 @MariaRivera_OC @DrPaolini Nope, not the same concept at all. A vaccine contains a small amount of biological agent, weakened or killed. It's small, but it's there, and it's enough to trigger an immune system response. A homeopathic "remedy" contains NO such ingredient. None. Zero. It's water.	-0.358	0.123
· @Fitz_RL @tcorrellrl yeah its quite unfortunate, his rotations were just getting good and he was evolving from a solo player into this. he learned so much from rlcs as well D;	-0.355	0.121
· MzDawnNicole Ok cool. Y'all the ones who finna miss the playoffs after starting 5-0. Ifs, woulda, coulda, & shoulda is y'all vocabulary consists of at this point. It's cool.	-0.352	0.118
· @niknik112 @nowthisnews It's a coin because it's shaped like one.	-0.352	0.118
· @Millard_Chochki @mtracey What's funny is that this scandal is rapidly turning into the #clintonrussia corruption. Humorous.	-0.35	0.117
· @roywlewis @benyc @RepMikeQuigley Your foul mouth is reflective of your hating heart. You're just another member of the Party of Hate #POH	-0.349	0.117
· @ljean @CBSNews Also that census is ESTIMATED. That means some marriages were older, some younger. And if you knew history and worked in genealogy as I do you'd know 16 was a typical age in many places in the 70s. Especially in the south.	-0.349	0.117
· Unpopular opinion for today: exo and bigbang is the best promoter for bts. Both groups literally gave bts career or they're going to stay unknown in Korea	-0.348	0.116
· @PopCrave @Louis_Tomlinson Everybody from one direction is better	-0.345	0.114
· @PamelaHurd11 @colettey6 Pam you're hysterical thanks for the laughs. Your very bad at this.	-0.343	0.113

· Josh Richardson is a better version of Klay Thompson Justise Winslow is prime Andre Iguodala Bam Adebayo is a much better version of Draymond Green mixed with Kevin Durant	-0.342	0.112
· @usarouse @kerrc17 @FOXSoccer @HirvingLozano70 @cpulisic_10 @ussoccer @miseleccionmx Hahaha keep riding his dick bro! Pretty pathetic you're so high on an unproven player. Yes he is u or	-0.341	0.111

POSITIVE: Dimension 4	coord	ctr
· We're going undefeated! #BYU https://t.co/XeyK6ExhRk	0.622	0.396
· I don't know if I s/b proud or disappointed that 5 people blocked me on my thread today. It would be great if people could refute things they had disagreement on with facts. It seems to be a high bar on Twitter. I appreciate @Boxerworks grace in recognizing a Twitter amateur. □ https://twitter.com/Boxerworks/status/956998687464488962	0.551	0.311
· I'm hoping 10 months from now trump will be #OutOfOffice #ImpeachTrump https://t.co/v5hEvZWvnY	0.522	0.279
· I think it's mostly because females are so weak, so it feels as painful as a heart attack to them. For men, it would just be closer to stubbing a toe.	0.514	0.27
· lol you're too westernized to even know what I'm talking about. https://t.co/gnlQeyl7Vh	0.497	0.253
· They know what's best for flyover country tho https://t.co/LrP59JbqWI	0.493	0.249
· If the #cowboys need me for locker room bulletin board material they are worse than I ever imagined. And nobody would know who Cole Beasley @Bease11 is if he played for anybody other than Dallas. @SportsRadioWIP	0.489	0.245
· .@RepMoBrooks: "Roy Moore will vote right, that's why i'm voting for Roy Moore." #ALSEN	0.465	0.222
· I will NOT apologize for being white. https://t.co/u5yZa5lkyS	0.463	0.219
· I have to say that I think if Andrew Breitbart were alive today I think he'd be extremely proud of the way his legacy of constant shameless lying has been carried forward. https://twitter.com/dandrezner/status/946133239680458752	0.458	0.215
· Sometimes I delete them, sometimes I can't help myself □ https://t.co/xovvflwD1e	0.451	0.209
· I think @JDERON21 "JV" team needs one https://t.co/ePqGuWVuSl	0.449	0.206

· No. We are not. #NeverBernie https://t.co/tJWy5dw5xx	0.448	0.205
· We can update you tomorrow night when you're in comms box with @Bazmccullum and I @KP24 #BBL07 @tensporttv https://t.co/MoYzTgqaOf	0.438	0.196
· That he's not Jon Gruden https://t.co/4Q7SBtTvMv	0.432	0.191
· Tonight's pin on @11thHour #MSNBC was inspired by a segment on @allinwithchris tonight about Trump playing golf 1 of every 4 days he's been Pres. WH trying to hide this. If the Mueller investigation is allowed to continue, maybe it'll give 45 more time for his favorite activity. https://pbs.twimg.com/media/DSGXnSwV4AE2OUg.jpg	0.431	0.19
· Could absolutely be me https://twitter.com/5HeadShawty/status/946224558515859457	0.42	0.18
· @Michael62380749 @chuckwoolery Nope. But I'll argue that it is. It's about all of us.	0.417	0.178
· Carson Wentz is fantastic. I'd take him over any other under-30 QB.	0.415	0.176
· I love triggering liberals. It's so fun and I think I did exactly that with this tweet. The debate is so entertaining to read, though. https://twitter.com/KBrocking/status/957275468369145858	0.412	0.174
· Dear Woodstock. @michaelcaldwell posts his votes. I've read them. You can do better. #VoteBlue2018 https://twitter.com/michaelcaldwell/status/965725070789431296	0.411	0.172
· Not long until it's decided! If you haven't seen this yet, please do check it out! https://t.co/jTS3dE2R7S	0.41	0.172
· .@NYCMayor wants to address #climatechange – but not if it means giving up his SUVs and helicopter rides. #ExxonKnew http://bit.ly/2ATdAvD	0.405	0.168
· I don't know why @SkyPelham is mad at me for slipping right before the camera timer went off. I had to hold her for balance. https://pbs.twimg.com/media/DUlAJ3KWkAAiw2G.jpg	0.405	0.168
· Trade Jayson Tatum I'm sick of him	0.404	0.167
· Steph Curry really couldn't see that kid of hear him. He was really far away. I hate how the media and folks on twitter are acting like that boy was easy to see. He's in a PACKED STADIUM!	0.401	0.165
· @MschRn @GoAngelo @seanhannity Because it's fun to bash @mmfa	0.401	0.164
· □□ I heard that I am GETTING a new POSTMAN next week, who is a SNOWFLAKE, Panty-Fa, Demorrhoid, Gun HATER...□ So I WHIPPED up this beauty JUST for HIM !! Woohoo !! □□ Silly WABBIT Soy Boy !! □□□ #MolonLabe #GunsLivesMatter #2A https://pbs.twimg.com/media/DWQfyv7VwAAIkz2.jpg	0.394	0.159

· Susan Sarandon is... - not a feminist - anti-Obama - unbothered by Hollywood sexual harassment - and still convinced Hillary Clinton would have been more dangerous than Trump My god, she's Phyllis Schlaffley! https://t.co/OtF1rvmB6X	0.394	0.159
· .@FoxNews is MUCH more important in the United States than CNN, but outside of the U.S., CNN International is still a major source of (Fake) news, and they represent our Nation to the WORLD very poorly. The outside world does not see the truth from them!	0.392	0.157
· @thebuddhacat1 @Resist_chick1 @realDonaldTrump That's right he tried but he fail https://t.co/9i2eC8hQdN	0.391	0.156
· Jobs you would never be able to do because they would require you to think of other people and then care about them, and then act on that. You're the worst. https://twitter.com/realDonaldTrump/status/946156544927977477	0.388	0.154
· Not that I'm Red Auerbach...but it's pretty clear to me that the Big 12 and the SEC are definitely the 2 best conferences this year at least in College Basketball....	0.388	0.154
· STOP telling me they're not messing with & MANIPULATING the #Weather unless your prepared to put up #FACTS! They have been playing Science #Weather Gods for over 100 YEARS! It's time to CLOSE DOWN THE SCIENCE LAB called Mother Earth #STOP THE #SPRAYING #WEDONOTCONSENT https://twitter.com/i/web/status/942528135853047808	0.387	0.153
· #MLKDay2018 While Liberals & #FakeNews spin @POTUS' words & call him a RACIST, I thank GOD that we FINALLY have a President that's improving the lives of ALL Americans of ALL colors. Obama & his minions were interested in OWNING votes, NOT truly helping people! #FridayFeeling https://pbs.twimg.com/media/DTX7XVVVAAAYalR.jpg	0.387	0.153
· I think ballot tracking is silly so....yeah. You're close. https://t.co/Pyn7HP0ovk	0.387	0.153
· I find this fascinating. The Broncos have won 3 super bowls in the same amount of time its taken the Chiefs to win one playoff game. I'm guessing it won't really register with Broncos fans when the Chiefs lose in the playoffs again. https://twitter.com/ArrowheadPride/status/946129073973157888	0.386	0.153
· Because many of us are #Labour supporter's but definitely not #Marxist #Marxism that's the difference can't won't vote for #JeremyCorbyn https://twitter.com/silverrich39/status/968582664776273921 …	0.384	0.151

· There is something really nauseating about this little clip..these kids think they are hot stuff; they do not realize that they are tools...flavor of the week...shills. Once CNN is done with them they will have to stay relevant in some way...now THAT is scary to me.□□□ https://twitter.com/HangEmHigh007/status/968288637421957121	0.383	0.15
· If I see anyone Open Carry a gun, I'm pepper spraying them immediately, taking their gun and holding him there till the police arrive.	0.379	0.147
· Richie told us Gophers were taking ownership. Tonight, it appears to be by shouting, "We're loud, we're proud &; we're horsesh-t."	0.379	0.147
· You're too flabby for me, I like my men fit, and smart. Sorry. https://t.co/rPbX6F6vy9	0.379	0.147
· @moldy_snowballs I feel sorry for you, it seems that you'll fall for any type of non-biblically-based mythology (that's so common place on the internet): https://www.gotquestions.org/Lillith.html	0.377	0.146
· I don't dislike DC or Zack Snyder. I'm just still waiting for him to make a good movie. https://t.co/3dEOPd2gif	0.376	0.145
· Can't believe they are not climbing higher to get #Mackiewicz @czapkins - he is the one injured #snowblind & #frostbitten - his blood is on everyone's hands celebrating @ZabRevol return https://www.theguardian.com/world/2018/jan/28/nanga-parbat-climbers-rescue-french-woman … #RIP #sendingboystodoamansjob	0.376	0.145
· Way to go everyone who likes eating these and thinking these are delicious... Now we can't have clean clothes □ https://pbs.twimg.com/media/DUmA_ShVMAEeUAG.jpg	0.375	0.144
· Oh yeah, I forgot. You can't blame a black guy of something even if he did it...because he's black. https://twitter.com/2the_hill/status/968531409571930112	0.375	0.144
· .@SenStabenow wants to fund dangerous sanctuary cities — cities that threaten public safety. It's time to #EndSanctuaryCities. https://t.co/VaaOV6SaA6	0.375	0.144
· youtu.be/_iY9Q_eWO8I @LordVinci_ last time y'all faced us in basketball □. I had forgotten about this one... See y'all been taking L's from us for a min fam □.	0.369	0.139
· @Monivader @aaronjhill @JacobAWohl @civilwartrust I know, 'it's so hard to except the truth when it's staring you in the face'	0.365	0.137
NEGATIVE: Dimension 4		
· @BradfordNims @PaulsEgo Why can't you and all the pro gunners just acknowledge the reality of your position? You don't care about people being gunned down	-0.537	0.295

in churches, schools, and other public places. You think your personal "Freedom" to have an AR15 trumps the rights of the others to live. Be a man.		
· @DonSather2 @AustenLied @WelterPeggy @andrewcockerpoo @peterdaou Hey bro, what voters were purged? Hillary's? Burnie's? Trumps? Yours? Sit down little man.	-0.449	0.206
· @SenWarren @LCVoters so why did you vote for the bloated, unaudited military budget - 1/4 of which could have funded repair of the all the problems in this country - homelessness, student debt, free college tuition, safe food, safe water, safe air to breath. You bloviate all the time. End the wars?	-0.428	0.187
· @BorchidJoseph @BrentBozell ...And your EVIDENCE IS??? The "word" of ppl who WORK FOR demorats-& who WAITED 38 YEARS TO UTTER A WORD?? LOL. NOT evidence--VERY LATE HEARSAY BS!!	-0.415	0.177
· @VictoriaBanvil2 @DavidBegnaud Right here and now...? Is PROMESA signed by Obama, which gave Puerto Rico an undemocratic US colonial junta. You think you can forget the murders, experiments, abuses, exploitation of lands for 119 years. You can't whitewash the shameful history of US colonialism...	-0.414	0.175
· @TRex86366601 Oh you so smart, huh? Fine... Describe the the Lorentz factor and how it relates to Einstein's Special Theory of Relativity. No Google...	-0.413	0.175
· @Charbrevolution @rolandbouman @Cattlechildren @BigBaldDr @EzzyMix2 @NiamhNolan8 @Takethatdoctors Seven years watching YouTube conspiracy vids and all the links and more. Meanwhile one child, not damaged, just different is where in your calculations?	-0.411	0.173
· @stonyjbc @Mattel @SuperChick356 Wait what? How did a Barbie doll trigger you all the way to accusations of opposition?	-0.406	0.168
· @usarocks_c @GregWest_HALOJM @nytimes How many close black friends do you have? How many PoC, especially women, are your peers or superiors in the workplace? How many sit on your city council, your school board, your party's committee, and your church committees? What % of incarcerated are PoC in your state?	-0.402	0.165
· @pmbasse @Lala_roo_boo @libertyladyusa @realDonaldTrump @Patterson4TX Who cares? No one! We literally started a bullshit war to take part of Mexico and we did. Shut up. You are not the only state in the Union.	-0.399	0.163
· @MrsAmy47 @tedlieu @SenateDems @SenateGOP @HouseDemocrats @HouseGOP	-0.398	0.162

@OfficeGovEthics based on what? And anonymous lawsuit filed during the election? So much bs		
· @PressSec Why are you posting this from your government account? What does this have to do with policy and the Trump Administration?	-0.383	0.15
· @davidhogg111 How much of THIS KIND of BULLYING did you do to Nickolas Cruz? Did you ever befriend or have lunch with ANY "outcasts" in your school? If not, then perhaps YOU should look to yourself for "solutions" --- Maybe YOU could have saved your classmates, by simple acts of kindness.	-0.378	0.146
· @NembuKol @mrpaluvets You dislike for RSS is patent & obvious. What do you expect in return ? Pls expect only 1 thing, One LTTE Tamil dead is One LTTE Tamil less. This is wanted by RSS People.	-0.373	0.142
· Hahaha only 65.000 H1-B visas are given a year. 1 million immigrants entered the country last year. What have you proved? These select immigrants must have a bachelors, lined up work, and are selected for science, engineering management positions.	-0.369	0.139
· @giddieupbitches @Early__May @TaylorEdwards99 @AJDelgado13 The point Your head Do try to keep up. Have you read the book? How can someone be brainwashed for calling attention to the fact AJ hasn't read it?	-0.366	0.137
· @RoryGilligan1 @FiveRights @daniellefreda Because one photo represents the entire dreamer population. Choke on your racist crap and die.	-0.364	0.136
· @JackSun01 @Trumpism_45 @realDonaldTrump @POTUS So if you were illegally spied on and persecuted for over 2 years with no evidence against you, while the MSM smears your good name on a daily basis, you'd just sit quiet and never speak up for yourself? Why would anyone stick up for you when you won't even stick up for yourself?	-0.362	0.134
· @SteveSchmidtSES Are you not familiar with deferments for college students? Trump stoped a murder in 91 how many have you stopped?	-0.361	0.133
· @intheMinorityOR @RonWyden @POTUS Poor Trumpflake does the truth hurt your witty bitty brain cell so much ch that the only argument you have is to toss out the word Liberal just because you have nothing else? Just shows your stupidity.	-0.361	0.133
· @Zenophile @markhentz @EllsBellsInPA Oh, OK, in a two minute peruse of your feed it looks like this is what you use when confronted with facts and being called on your nonsense. The irony is that the post demonstrates exactly the type of hideous morality that people outside the JoePa bubble view as nauseating.	-0.361	0.134

· @tyrasquared @tamralee68 @zachkap15 @MrSmithtim @Inafume @DanielGdh12003 @DavidCornDC @FoxNews Don't worry AI isn't that good yet. You have to say machine because plenty were done in paper form, digital votes have way more vulnerabilities.	-0.36	0.133
· @SassySculptor @Success87473781 @ObamaMalik Michelle Obama went to Princeton and Harvard, Melania got her jumpstart from doing soft-core porn photo shoots. Take a seat you inbred fuck.	-0.359	0.132
· @issyelliot @Unite2020 Why stop with the wall? Why not fund ALL government programs that way? Planned Parenthood, NPR, veterans benefits, etc, etc,.....	-0.357	0.131
· @KatyDidlt @GenMhayden @jmclaughlinSAIS You claim you want to ban a whole religion in your bio. Kindly STFU about Trump's lies being outed as somehow representing an "assault" on his 1st Amendment rights. You are fundamentally against American ideals.	-0.356	0.13
· @_MrsAtheist_ @JJohnsonLaw @VanityFair How about you get over a small, innocuous joke and do something productive?	-0.356	0.13
· @atkins_brock @TraderFrog @mtbrown94 @pgdaly84 @JDoc_son @TheCoreyColeman And what you won a couple Mountain West championships? Against who? Boise State? □□□	-0.355	0.129
· @WJ_Armstrong @canokar @06JAnk Well then you should never complain if you ever get silenced or censored for simply speaking out, w/o any harassment, insults, etc. What you waiting for? Block me already!!	-0.35	0.125
· @old_warrior1 @csutton1959 @4WCowgirl @AdamSchiffCA What is your goal with this response? Do you ever just defend your positions? Like the validity of CBO scoring?	-0.35	0.125
· @IRdotnet @realDonaldTrump @jamiejmcintyre @dcexaminer What are your thoughts on losing to Obama again in the Gallup “most admired male” award?	-0.348	0.124
· @leenathanael @RepWilson @subverzo You call it grandstanding when she, by name, commends Republicans for all they did to expedite the naming of the building? Do you understand what cooperation & consideration is? Take off your glasses, they are clearly the wrong prescription.	-0.347	0.123
· @oewonah @CallmeAlfredo @ethereal_17 Bigger is no pride ...a lotta human beings grasping for everyday amenities?.......don't pride yourself for just being big...	-0.347	0.123
· @slushy___ @Education4Libs wow ... you're all about the unfounded accusations. What has TheMister said that assigns an ideology to you?	-0.347	0.123
· @Lynxos1971 @tedtully @BrexitTory_ Grow up you crank! You live in a city destroyed by a socialist council experiment in the 80's□	-0.345	0.122

· @420Lisa No will not go to jail! Your are missing the real point listening to fake news. Why don't you complain about Amazon paying no federal taxes?	-0.344	0.121
· @JulianAssange @laurilove hey julian, you did something good for once. helping out and bringing attention to Lauri Love. Well done mate,	-0.343	0.12
· @broncossuk @thekidmcmanus @Patriots @Broncos Sure, you do. You continue to lie, which is your specialty. Nobody in college, and nobody who has a job, can be this much of a loser. Again, get out of your mom's basement, and get a job.	-0.339	0.118
· @breton_anne @WokeFormerLib @Jonessmithadams @chidori69 @jackschofield @ejethan123 @realDonaldTrump @NancyPelosi @POTUS You have proof of that? So far there has been no proof that President Trump did anything wrong. There is plenty of evidence that members of the previous administration tried to stage a coup against our President though. Also, if you are in Canada, why do you care?	-0.337	0.116
· @LisaFromEarth @washingtonpost And what does your tax return show. How much do you get back????	-0.335	0.115
· @jaketapper @KurtSchlichter When will you revisit Juanita Broderick's @atensnut story? gvien all the recent developments and sudden media interest in decades old allegations.	-0.335	0.115
· @THEOrangeDog Coming from someone who doesn't have any idea what negative publicity does to coaching searches and recruiting	-0.334	0.114
· @iMthinkingPinoy Ikaw, who did you sell to your dark soul to?	-0.333	0.114
· @JacobAWohl @realDonaldTrump Obama gave away tax payer dollars to big banks and auto companies. What did the America people get in return? Their jobs were shipped to Mexico thanks to NAFTA	-0.332	0.113
· @thatcameraguyc1 @FoxNews Is all of this clear to you now, chico? Or do you lack the reading comprehension skills to follow? Hmmmm	-0.332	0.113
· @roofer_fl @Vocelle733 @TomiLahren Do you have any proof that purple unicorns do not exist? Same thing, cunt.	-0.331	0.112
· @gregdeckard @misterbumface @FoxNews @mattgaetz @realDonaldTrump Why do you lie. 4 indictments, 2 for money laundering and 2 for lying. Not a single indictment has language pertaining to a conspiracy to collude with Russia. Guess what? Dt is potus. Keep hoping and dreaming that Mueller recommends something congress would accept	-0.33	0.111
· @mark_earnest As usual, just in your pathetic little bubble, do you know what words " robust economy,	-0.33	0.111

created millions of jobs, and the lowest unemployment" mean?		
· @SueCFlorida @damartin32 @LizCrokin @Avonsalez @buildthewall_20 @emarie1225 @LemboPhil @MScipio_African @t193931 @TanHaley @Tspinnerchaser We cannot begin to talk about shadow corruption when blatant disregard for the rule of law is taking place in place sight and the potus you all are saying is here to fix all of said corruption is not lifting a finger to clean his own house. Explain, anyone?	-0.328	0.11
· @Comey You lied to congress, failed to do your job and protected a criminal. You don't get to talk about freedom of speech any more.	-0.328	0.11
· @kristiemacris @HarrietNix @dukkiller9 @StutzmanAnn @LANURSE1 @kar_bear77 @artist35 @chic_savage @wolfgangfaustX @nikdpik @AnneMarieThiel @thumperalpha @MDDeplorable @ThinBlueLR @sgmills52 @PoliticallyRYT @steph93065 @plagueoflegions @hatedtruthpig77 @Education4Libs @GaysForTrumpFL @Justd1989 @realDonaldTrump @POTUS @FLOTUS Huh?? National-so you want single payer-do Canada, do England, do socialism. You must have embraced Bernie on his freedom mirage-all free and no one pays...doesn't work anywhere	-0.327	0.109

POSITIVE: Dimension 5	coord	ctr
· Do you know what a question mark is, stupid? How's this for a new tactic: You are a brainless trump cult worshipper and yo mama is a bitcoin dumpster ho. https://twitter.com/SpecialEDxx/status/934095239890976768 …	0.6	0.426
· Why isn't @brianstelter, whose beat is covering the Media, reporting on this? Hate me all you want, that's fine, but is this OK, Brian?	0.552	0.36
· So what should we do @AnaKasparian start a war with the Philippines? This is your problem. Mind your fucking business. #TYTLive	0.548	0.356
· @USAF_Frye @kmassey04 @TomiLahren @foxandfriends That's why your a single trump supporter... and I'm not fake my followers know what I really look like get it through your thick skull because I'm only saying this once more this is a stan account.... pic.twitter.com/QMKqPwRxbQ	0.519	0.319
· @LexiHunt00 @mfleming877 @cathy58444301 @adarondax1 @Edible3Ball @realDonaldTrump @caramastrey @flowers3712 @SandraJH13 @cindyknoxville @marklevinshow @wishgrantlotus	0.487	0.281

@CalebEatsBacon @California4Trmp @POTUS @MissTigerAngel @Greggorj @Cyptocon70 @BrutalVeracity @jaydixson Report what?. You don't like it.??? Then leave Twitter. This is our place now. Go back to Facebook with the rest of your kind. Buh Bye!!! @BarackObama @POTUS44 @MichelleObama @ObamaFoundation https://pbs.twimg.com/media/DXG4bYaW0AAehCM.jpg		
· Proud to work for @POTUS who stands strong for Israel and all the Jewish people. #NeverForget #NeverAgain https://m.youtube.com/watch?v=wHqcvxBCm2A …	0.486	0.28
· Curious @HawleyMO what's your favorite sidearm? Or rifle? Do you know much about the weapons you're calling to control? #2A	0.483	0.276
· "Where are you from?" "Sarawak." "Where's that?" "The place where Anaconda:The Hunt for The Blood Orchid was filmed." "OH I LOVE THAT MOVIE............ so................ is it rea—" "NO." https://twitter.com/ShuckMyBhauls/status/968680694112071680	0.477	0.269
· @Andyxfish @trinafoxxmovie @Alyssa_Milano Umm...you bio mentions nothing of this and how is it you are able to be on Twitter? □Maybe you should be doing something more productive?	0.475	0.267
· Bloop! @DohertyShannen called out sister witch @H_Combs as being totally negative about the reboot & that she don't want none of that in her life!! She wants us to give the reboot a chance! Which #Charmed star's side are you on?! #CharmedReboot	0.471	0.263
· @JBassett85 @colinJclayton @nottjmiller I'm the bad guy!?!? You know what dude YOUR what is wrong with this great country the US of A. We as in me and TJ don't need any of you soft charmin style baby sh!t wipes saying nasty comments. Thanks but no thanks J. Your a shitty photographer. Good luck. Now speak to "the" □□	0.465	0.256
· I know that the the NRA owns you @marcorubio @realDonaldTrump @FLGovScott but I mean common why is it so hard for you guys to rip off the shock collar the NRA has on you because at this point y'all are like a bunch of really stupid sharks that thing you have power.	0.462	0.253
· @TravisAaronWade @Eminem I don't think you understand what NC-17 means, but... So you're complaining about conventions with explicit behavior warnings and inappropriate behavior at those conventions, then post what you consider mature content with a warning to the same crowd?	0.459	0.25
· How do you rate @realDonaldTrump 's #progress in his 1st year as @POTUS? Please vote & RT Ì†ΩÌ∏ä	0.449	0.239

· @BeerMeMarge @KeepTXTX @KellyGirl2018 @gsteck74 @ArleenCandiott1 @RickCChambers2 @Steve_Pippin @Larry_in_Ohio @azteetime @cCoVrHlt @hogmania2 @Jasmine8137488 @JsTacoma @LaunaSallai @Maggieb1B @PatriotJeweler @Smartassy4now @soulrelevant2U @Susan_Texan @time_kelly42 @tld5678 @TSpriDeplorable @DoubleTMisterB @zack_nola @POTUS Do... do you think that an anecdote and anecdotal evidence are the same thing? Damn, you might want to delete that tweet. It makes you look real dumb: https://en.wikipedia.org/wiki/Anecdotal_evidence …	0.445	0.234
· @James12jl @Cella5212 You know what separates a top 10 yardage ranking over where the Packers are right now? 20 yards.. lmao □ that's it.. so to think that's some kind of indictment is amateur lol	0.444	0.233
· SixTEEN HOURS! 16 hours...where is my account access? I want my account back! ABSURD and BIZARRE!! I'm running out of patience. @SeanHannity #PelosiSHUTUP #buildthewall #draintheswamp #rosap #releasethememo	0.442	0.231
· @bloodyhell_rblx @dilekaksu @mattbramanti @PeonRevolt @davidfrum lol...Hitler made it against the law for Jews to own guns. And guess what happened after that...YOU are a holocaust denier. stop being anti-semitic!!	0.437	0.226
· @PenlandKW @SenSchumer @CFPB Hey idiot, the President picks the Director of the @CFPB that's how Congress set it up. Learn to read. You apparently are clueless about how dictators rule. Educate yourself.	0.436	0.224
· @SLandinSoCal @realDonaldTrump #OperationMockingbird #JFKFiles Released files revealed at the CIA was controlling the media to generate propaganda. Don't believe what you hear on mainstream media.	0.434	0.223
· @AnthonyUSAA @realDonaldTrump Do you know what Russia wants? For the Taliban to create chaos in the region. Why do you continue to cater to Russia?	0.431	0.219
· lol you're too westernized to even know what I'm talking about. https://t.co/gnlQeyl7Vh	0.429	0.217
· OPINION: Hey, @NFL owners, tired of half-empty stadiums? This isn't hard. Tell your players to stand up for America https://t.co/yLWvXRXdwI	0.428	0.217
· @KinraIce @yourfaveclete @MCL1381Bones @JacksonLeeTX18 Again. No FACTS. Do u know what facts are? Or what evidence is?? □♀	0.426	0.214
· @Resist_chick1 @thebuddhacat1 @The_UnSilent_ @realDonaldTrump Then how come you don't know that the Electoral College is the one that selects the president?	0.424	0.213
· @Cernovich Is that why you use your “Alpha male” self and duck out from people calling you on your bullshit	0.422	0.211

Mike? Because you can't fool people who actually think? You're a shitbag waste of cum. See you in NY. https://pbs.twimg.com/media/DSHGrnlXUAAIVk6.jpg		
· When you forget (or never even knew?) that Noakes is also a qualified MD..... □ https://t.co/yG4SdhZoHz	0.421	0.21
· What going on ? @AITCofficial will raise this in Parliament. Bad enough you cut rates, worse still when the House is in session https://twitter.com/Truth_of_WB/status/946217798182739968	0.407	0.196
· @Dolphieness @GenMhayden @jmclaughlinSAIS You're false equivalencies & general disrespect for what actually makes #America great let me know that you're either a) an idiot b) a #RussianBot c) a #TreasonWeasel d) all of the above	0.405	0.194
· Reason #33 why @MildlyAmused is awesome. https://t.co/9yhL6Tj07i	0.403	0.192
· @HunterFallacy Aw, I see what this is about. You're upset that I didn't remember you. That's so pathetic. Sorry, doll, but obviously you didn't stand out from the crowd of all the others in that thread. You poor thing. I hurt your feelings and for that.. well, you are just sad. #GetWellSoon https://pbs.twimg.com/media/DUnm86IVwAE7V6V.jpg	0.403	0.192
· Copeland preaches a different Gospel...a different God than the one in the Bible, and THAT has lead to his racks-that's my issue tbh" can u c how this reply leds a person to think ur saying its ok for another to get rich off church	0.401	0.19
· @imkeshav @AndrejGregus @BitcoinCashD And you were the lucky one to get one of the two !! Your response to measure someone based on # of tweets shows how immature you are similar to you trying to predict the price movements for a crypto based on technical past data/movements ! Best Of Luck !	0.401	0.19
· @NoBongo @steffan_nancy @trevorHartje @realDonaldTrump You do understand how population centers work right? Where ya know, the people are?	0.394	0.184
· Why are you so afraid to pay your child support? Afraid it will eat into your hair gel and musket budget? https://t.co/RZduYvNAcE	0.392	0.182
· @RealJamesWoods And your point is? So when I'm walking through a mall full of Europeans should I wonder why they're not in Europe? Why are all these white folks walking around? This is America where people of all ethnic and religious backgrounds are welcome regardless of what any of you say.	0.391	0.181
· @JosephSommers4 @Cernovich Karma is a bitch and so is your bitchboy, the triggered snowflake Mike Thernovich. That's who you worship at the altar of, a weak	0.391	0.18

beta whining that some soft libs are mocking him? What weak, sad non-men you all are.		
· @MarquisHorace @thehill Because you don't know what you are talking about, you are just a parrot that repeats what is on MSLSD.	0.389	0.179
· @Onawhim What tantrum? I am laughing at this movie tanking, people finally realizing how unethical journalism is and you proving my point of who the real mysogynists are. You think she needed rescued. Therefore you think she is weaker and dumber than you. Pathetic.	0.387	0.177
· What game next year are y'all gonna hang up the banner for your ACC championship? https://t.co/0AtL43rAv6	0.387	0.178
· Remember, everyone loves wine in their stocking this holiday season. Check out the great @trumpwinery selections this #BlackFiday weekend!https://pbs.twimg.com/media/DPewpgiVAAEeh7Y.jpg	0.38	0.171
· I been studying this whole flat earth vs globe thing... and I think I may be with Kyrie on this... b4 you judge do some HW but what do you guys think?	0.38	0.171
· @thowelliv @poseidon664 @justin_fenton @baltimoresun Impressive... “circle jerk over police”... Did those kids you are showing off teach you to speak that way or did you come up with that quote all by yourself?	0.378	0.169
· @AlwaysThinkHow @huckfinn22 @peterdaou No, you are divisive and that's exactly what you're trying to do. All of your rhetorical nonsense is just lies. Do you even know why you don't like him?	0.376	0.168
· @Maurice22015378 @miss_speech @realDonaldTrump Natural English? Lol... Because English is so original... And you're about as hilarious as cancer... I find it amusing oh how funny you think you are	0.375	0.166
· @Jdcruz1977 @GregWest_HALOJM @FOconflict Yes, you can, Jason, proving what a traitorous little shit you are. That your treason is driven by racism is appalling beyond words. Go play in traffic, shitstain...	0.374	0.165
· @PattyArquette @VanityFair That's nice. Now @VanityFair knows how deranged Clintonite loyalists are, and what us dull normals have had to deal with.	0.372	0.163
· If the #cowboys need me for locker room bulletin board material they are worse than I ever imagined. And nobody would know who Cole Beasley @Bease11 is if he played for anybody other than Dallas. @SportsRadioWIP	0.37	0.162
· Make this go viral. #QAnon made a threatening post the other day about "BRIDGE" and look what I said it meant (according to our insider)... and look at #Arlington. #QAnonsaTerrorist #qanon #anthrax https://twitter.com/JoabsInHisArmy/status/968088597588037632	0.368	0.161

· Can't believe they are not climbing higher to get #Mackiewicz @czapkins - he is the one injured #snowblind & #frostbitten - his blood is on everyone's hands celebrating @ZabRevol return https://www.theguardian.com/world/2018/jan/28/nanga-parbat-climbers-rescue-french-woman … #RIP #sendingboystodoamansjob	0.367	0.159
NEGATIVE: Dimension 5		
· @Loc8YourDignity @cookiebaker57 @thehill & "serviced" v "survived" but I got your meaning. Of course he has emotional knowledge & clearly he's been well coached on gun control ideology, when he goes off script he can get into the weeds. He's a 17 yo kid, no matter his o/ward maturity, he's still developing cognitively.	-0.432	0.221
· @SuperSteveDV @SthrnMomNGram @karinagw @HebrewNational You have no prof he did it and don't even want to give him a chance. If there is prof then he deserves to be punished but until such time he remains innocent	-0.424	0.213
· @floweraldehyde On the contrary, Apart from my school education which totalled to around 2 lakh from Kindergarten to Intermediate, I didnt receive any money for UG (around 1 cr paid by taxpayers). My PG also will be paid by taxpayer (most probably) around 2 crore rupees.	-0.422	0.21
· @shannonrwatts Wealthy white Bernie supporters are obnoxious. Like Sarandon, they have other places to live and the means to do so, unlike those of us stuck here under Trump. HRC voters would've voted Sanders, we were not shown the same solidarity.	-0.421	0.21
· @that_nocoiner @indystar We have no moral high ground here and we missed an opportunity to have an important and necessary conversation. Instead, folks like you can feel vindicated for not liking "those people" because, well, they don't like us either. https://t.co/YK235qMAui	-0.412	0.201
· @bonnieotterson @DineshDSouza @realDonaldTrump 2. The 16 yr old, & Allred, had a fake yearbook. They will not let an independent handwriting expert look at it. Of the two remaining that said they dated him when they were just 17 & 18. One worked on Hillary Clinton's campaign and second one MAY have been paid. Typical hit-job!	-0.41	0.199
· @GregWest_HALOJM @NikkiProverbs31 @DutyOfAPatriot @Bobby_Axelrod2k Or, when it doesn't go so perfectly, homeschooling becomes the inbred among education options who can't count past 10, but has 16 fingers (ﾉﾟ－ﾟ)ﾉ Srsly, homeschooling is often used to keep kids in Bible study for 8 hours/day, once Ma realized she doesn't know fractions.	-0.389	0.179

· @DineshDSouza It's pretty obvious from the people opposing him that they're more concerned about no more exposure of the corruption they've hid so long	-0.386	0.176
· @SophiaHelwani @eddiebravo @BrendanSchaub @ChaelSonnen @MMARoasted @TimKennedyMMA @dc_mma @jeffwagenheim @MMMAjunkieGeorge @MMAjunkieJohn @MMAjunkieSteven @MMAjunkieMatt @Benaskren I've never faked any followers. I definitely have a lot of foreign bots offering sexy time in my mentions though. I think all popular accounts have a certain amount of fakes but I've never done anything to acquire them.	-0.385	0.175
· @RandomgirlLisaE @MJLCatholic @JamesMartinSJ the church will shrink, but she will never die. It's being purged of the nominal so-called Catholics. Also known as the politically correct, leftist, modernist, progressive types.	-0.383	0.174
· @RebelLucy_Roze @TurnTNBlue @Rightwingmadman @DonaldJTrumpJr @wikileaks Comprehension really isn't your strong suit. I never rescinded any statement. There is currently NO investigation on HRC. There have been 2 before and she was cleared both times. If they do a 3rd, I'm sure they'll end up with the same result.	-0.377	0.168
· @MHJulie @anthonyzenkus Broaddrick wasn't a "great witness" as to the obstruction of justice charge in the Paula Jones case since she had nothing to contribute to it in her testimony. But Starr and his team absolutely found her claim of being raped in 1978 entirely credible.	-0.375	0.166
· @NickVikingChild @MarkACollett I know and it doesnt make the quote any less true. Most women did work until they were married or after when they became widows. During marriage they would take the much more useful task of taking care of the family and household. Spinster was also the term for a unmarried woman.	-0.372	0.164
· @AnonymissEve @thecftwatchers @Baby___Del @cdwillett66 @HeliaERossini1 @jwhaifa @yceek @OurBudA @danaquinns @FrankDaTank2004 @chasegolf2069 @ScottRickhoff @theory7ix @abhishekrana981 @242cats @oldtropper75 @Cleverfun66 @BIGSEXYYT @DunbarJoseph @wolfgangfaustX @FoxNews @peterboykin @jeffsessions You failed to make a point, so you resorted to just tweeting an insult. If Trump's first year in office wasn't enough for me to regret my vote, his followers here at Twitter sure are. Bigger snowflakes than any SJW I've ever seen	-0.37	0.162
· @BroadStBulliz @realDonaldTrump Oh yeah I'm sure he was lying despite not knowing he was filmed and potentially risk losing his job. It's been a year, there's no evidence and if there was he would've been in jail a long	-0.368	0.16

time ago. Once the DOJ acts, it's all downhill and screaming to the sky for you bud		
· @jm95serendipity @Fatemeh_Cool im not worried :) aint nothing they'll do about my tweet sis shsjdjdj	-0.364	0.157
· @Dooothe @magdutza_fl @gillovnot @TwitterComms It has been already proven it was not Trump who started that birthed movement. Hillary started it when she ran against Obama lol Climate deniers? There is actually no such thing. 37,000 Scientists have disputed “Global Warming”, and have always advocated “Climate Cycles”..	-0.363	0.156
· @DutyOfAPatriot @NIVIsa4031 I'm a former trucking Co owner. Fleet of 42. If truckers wld run legal, not play with their log books, ILD wld not be needed. All my trucks had OBR/ILD since 2011. No problems. Drivers made $$$.	-0.362	0.155
· @Ajairah1 @johnahodgsonf1 @s_herdis @akalamusic Nar still fucked attitude to have Subsided rates of white men could be said is party down to being told there racist have privilege have life easy blah blah blah.	-0.362	0.155
· @theealig forced me into making a twitter and she ain't even active on here wuuuuteeevvvuuuuh	-0.36	0.154
· @BBCSportScot I'm neither a hibs or a hearts supporter, but Shaw was 6 yards out, he should never have hit the crossbar it was an easy Tap In hes got to blame himself	-0.355	0.149
· @vanOnselenP It's also not right to name the men! If they face an investigation then everyone should be anonymous until the official outcome is released. It's about being fair to all the parties concerned.	-0.354	0.149
· @lookawaygirl1 @davidhogg111 The adults have had 20 years and did nothing. They are still doing nothing. He is articulate, informed and giving America a lesson in democracy and the political system. If he were my kid I'd be beyond proud.	-0.352	0.147
· @ellievhall @xPulisic10 It's wrong they picked a women, obviously she's gonna be biased and side with the women, just goes to show women are just not fit to do anything besides making me a sandwich.	-0.351	0.146
· @Koorivlf @Zaush Satan and his fellow demons will one day be sent to the Lake of Fire, along with all unrepentant sinners. (Keyword: unrepentant.) Demons flee from individuals who call upon the name of Jesus and resist temptation. But their presence is in the air over every nation.	-0.347	0.143
· @harrygarris14 @Adam_Mares Everyone would have lost their shit if Plumlee played for Jokic. This wasn't on the coaches, they tagged a 19 point lead. It happens but it sucks. Denver has been one of the better teams at holding leads this year.	-0.344	0.14

· @IngrahamAngle BREAKING NEWS: Trump can get 70% of the Illegal Immigrant DACA vote by telling the Spanish community before the midterms if they vote Republican he will get DACA done. Trump keeps his promises, he has a proven track record, Democrats never keep their word. Trump can sell that!	-0.344	0.14
· @cjdtwit @Tori2uTori @unconcious0 @atillathehun3 @JebandCorey @mattlogical @stand4honor @lajohnson1959 @Seeds81Planting @uniquedeehan1 @saiberspace1 @CAoutcast @UnionMan16 @RoryGilligan1 @bronson69 @Alice00581238 @SKSSKanz @TechQn @RUSTIMCCOLLUM @RealQuietMouse @TabbyWesa @ShoreyMichael @Dianalo43311735 @LandenSmith11 @RobStLaurent3 @billbenedict61 @DandAExperts @ravena68 @Becky91663 @SenateMajLdr @SpeakerRyan @POTUS No! They actually hired two guys who were embroiled in election fraud during 2016 election to head up the election oversight committee.	-0.342	0.138
· @4everNeverTrump @DonaldJTrumpJr @wikileaks PROVE Russia gave the info. Show us irrefutable proof - not a bunch of heresay. It's been a year + & ZERO proof has been given. All Mueller has for his HUGE taxpayer LOSS is 2tax evasion chgs which had NOTHING to do w/ election. Awake Americans know it for DNC Hillary lie it is.	-0.341	0.138
· @pilaraymara @NicolaSturgeon "handing devolved powers back to them " Please provide an example of a devolved power, just 1, that will be handed "back" to WM. Note: devolved power must have been within the remit of Holyrood.	-0.34	0.137
· @B0ST0N_B0B @here_mel @comicAnton @MissGFYCuffy @gramV319 @truthtalk4once @nastypantsuit @moststylish1 @KruseKimberly @icklepepper @Chrisjo13267688 @MichelleResists @Victoria_Iamurr @humanvisionary @AkulakSINY @Grandma_Sheila @ellenc53 @Intel3210 @yesimpeachnow @MelodyT86535977 @lockedmith @bkgut3 @DeborahResists @fdell3 @JoeB2112 @DeViLhErSeLfIe @Nelle19711 @HallieShera @Milead7 @RiseUp4ALL @SarahShera13 @beeby0420 @Pitchinafit1 @IzJustMyOpinion @Unpersuaded112 @DJ_PsychGuy @NiceDreamWithU @OriginalKMF @galantelaura @reddit @YouTube Sanders' lack of outreach to minorities may have cost him the nomination, but Hillary wasn't exactly a favorite with minorities, either. Regardless, Sanders has voted consistently in favor of advancing civil rights and women's issues. I just think he's too old.	-0.338	0.135
· @MissButter @realDonaldTrump Mueller is over reaching his Authority! He was hired to investigate Russia	-0.338	0.135

Collusion with Trumps campaign and White House people..Not prior to 2016 when they were Not associated with Trump.		
· @Mel_Ankoly @dianelyssa No. She won the popular vote. DNC rigged primaries for Obama by stripping FL and MI of pledged delegates, while allowing TX to split its pledged delegates by holding both a primary and a caucus.	-0.334	0.132
· @nukemanW88 @NFLResearch @sportywineguy Everyone said he had no shot vs Atl, then even less against Minn.	-0.33	0.129
· @realDonaldTrump I'm surprised you took time away from golfing and watching cable news to do something else!	-0.328	0.127
· @MikeHeed @KJones34301414 That “part” should have never been said so keep smashing burning and boycotting we libs will be here for the entertainment every time	-0.328	0.128
· @queenideator @dd4iu @atlgrace2 @askjillian @FBI Eleven women came forward during the 2016 campaign to accuse the then-Republican presidential candidate of unwanted touching or kissing. Other women accused Trump of walking in on them when they were undressing at beauty pageants he owned.	-0.326	0.126
· @BRGooley @PattyArquette @YRabl @summerbrennan And yet he lost. Badly. Rejected. By. 4,000,000 votes.	-0.323	0.123
· @MikeOneshot @CNNPolitics Don't worry Trump might be picked some year but for now he's just a loser.	-0.322	0.122
· @iLLWiLLTHEMiCK @just8anapple @markcubsfan @lawyer4laws @Porkmason @sequinpants @piersmorgan Not according to some. Radical feminists argue, based on the flawed univariate analysis of gender disparity in salaries, that they should be paid the same irrespective of differences such as hours worked, performance, agreeableness, etc.	-0.321	0.122
· @RichieBFE @hollyhaygood @ChrisButcher83 @Bingchemtrails They stop and have already moved, proving a rotating earth. Name failed experiments. You have researched nothing I guarantee it. You don't understand science at all, I guarantee that as well. Try me, every flat earther that has, was proven wrong.. so please do so I can laugh.	-0.317	0.119
· @SteveP37 He's just as bad as Lampard for going to City, don't defend him. Both snakes and have lost their status at Chelsea	-0.316	0.118
· @TheSilentLOUD @KimChris80 @RNcat50 @alohabrianb @davidhogg111 I just loathe people who claim to love Jesus being vile to kids who have just had the most traumatic experience in their lives. I can't begin to comprehend the terror brought to them by Trump's	-0.314	0.117

gunman, and now they are getting more abuse from Trump's twitter trolls.		
· @RedheadAndRight @eugenegu @ThomasS26985726 I wouldn't ever trust a Dr who takes a stand against white people in turn, his Hippocratic oath!Absolutely disgusting! He should be fired!	-0.313	0.116
· Nevermind that they let Roy Moore run in 6 Elections but wait until a month a senate election to come out and one of them just happen to be already outed as a Hack for the Democrats but Nice Try □	-0.312	0.115
· @raslady1 @Imperator_Rex3 Yes, Trump was so poorly behaved, even as a child, that he had to be sent away for the sake of himself and his family. Trump continues to struggle to respect the rights of others, to accept that the rules apply to him, and that self discipline is a virtue. God point.	-0.312	0.115
· @DrPhil I would have taken that same hammer & whooped Brandi's ass if she did that to my stuff! You ain't gonna break my expensive shit & get away with it, no way! Woman or not, I'd beat that ass for breaking my items. She asked for it.	-0.311	0.115
· @Glock_Coma_ @FoxNews @foxandfriends Never said she wasn't. If she wants to participate in victimizing those without power to punish someone for her child's death, it's her right time do so.	-0.309	0.113
· @lolli_logan @24baseballReed @artsycarol @Evan_McMullin Here in Australia, people see Trump as an absolute joke. My American husband gets asked about him all the time because Aussies can't believe he was elected.	-0.309	0.113
· @michellesawyer6 @Greytdog @tonyposnanski I've been blocked by him for a year now, sent a hello tweet to Mrs Chachi yesterday, not very friendly lol https://pbs.twimg.com/media/DUlAmITUMAAMu5H.jpg	-0.308	0.112

References

"troll" NetLingo. (1995-2015). NetLingo® The Internet Dictionary. Retrieved on 31st December 2015 from http://www.netlingo.com/word/troll.php.

"troll" Oxford English Dictionary (2017) Retrieved on 6 September 2017 from https://en.oxforddictionaries.com/definition/troll.

@Cromwell606 (2019) https://twitter.com/Cromwell606/status/1087798163069652993

@youNEAR_UHgee (2016) https://twitter.com/youNEAR_UHgee/status/814546280106119169

Abril, E. P. (2018) Unmasking Trolls: Political Discussion on Twitter during the Parliamentary Election in Catalonia. *Trípodos* 39: 53—69.

Adami, E. (2018). Styling the self online: semiotic technologization in weblog publishing. *Social Semiotics* 28(5): 601—622.

Ajayi, T. M. (2018) The pragmatics of commendation in selected "abusive" utterances among Yoruba youths in Nigeria. *African Notes* 40 (1&2): 80—94.

Alkharashi, A. and Jose, J. (2018) Vandalism on Collaborative Web Communities: An Exploration of Editorial Behaviour in Wikipedia. In *Proceedings of the 5th Spanish Conference on Information Retrieval (CERI'18)*. Zaragoza, Spain (26-27 June 2018).

Alorainy, W., Burnap, P., Liu, H., and Williams, M. L. (2018) 'The enemy among us': Detecting hate speech with threats based othering language embeddings. arXiv preprint arXiv:1801.07495.

Anderson, C. (2018). People are trolling their moms with this text about microwaving turkeys and the responses are amazing. *BroBible*. Retrieved on 4th February 2019 from https://brobible.com/culture/article/how-long-microwave-turkey-memes/.

Anderson, C. (2019). How Twitter Needs to Change TED interview with Jack Dorsey. Retrieved on 01/07/2019 from https://www.ted.com/talks/jack_dorsey_how_twitter_needs_to_change/transcript?language=en.

Ansong, E. D., Takyi, T., Damoah, D., Ampomah, E. A., and Larkotey, W. (2013) Internet trolling in Ghana. *International Journal of Emerging Science and Engineering 2*(1): 42—43.

Arancibia, M. C. and Montecino, L. (2017) The construction of anger in comments on the public behavior of members of the social elite in Chile. *Discourse & Society* 28(6): 595—613.

Ard, J. (1982) Auxiliary *do*: support or emphasis? *Linguistics* 20(5-6): 445—466.

Asenas, J. J. and Hubble, B. (2018) Trolling free speech rallies: Social media practices and the (un)democratic spectacle of dissent. *The Journal of Culture and Education* 17(2): 36—53.

Asención-Delaney, Y. (2014). A Multi-Dimensional analysis of advanced written L2 Spanish. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi- dimensional analysis, 25 years on: a tribute to Douglas Biber,* pp. 35-79. Amsterdam: John Benjamins Publishing.

Atkinson, D. (1999). *Scientific Discourse in Sociohistorical context: The Philiosophical Transactions of the Royal Society of London, 1675-1975.* Mahwah, NJ: L. Erlbaum Associates.

Bacile, T. J., Wolter, J. S., Allen, A. M. and Xu, P. (2018). The Effects of Online Incivility and Consumer-to-Consumer Interactional Justice on Complainants, Observers and Service Providers During Social Media Service Recovery. *Journal of Interactive Marketing* 44: 60—81.

Badjatiya, P. Gupta, S., Gupta, M. and Varma,V. (2017) Deep learning for hate peech detection in tweets. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 759–760. Perth, Australia (April 3-7, 2017).

Baker, P. (2001) Moral panic and alternative identity construction in Usenet. *Journal of Computer-Mediated Communication* 7(1). Retrieved on 6th August 2019 from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.2001.tb00136.x

Balci, K. and Salah, A. A. (2015) Automatic analysis and identification of verbal aggression and abusive behaviors for online social games. *Computers in Human Behavior* 53:517—526.

Balusu, M. R. B. Merghani, T. and Eisenstein, J. (2018). Stylistic variation in social media part-of-speech tagging. In *Proceedings of the 2nd Workshop on Stylistic Variation*, pp. 11—19. New Orleans, Louisiana (June 5, 2018).

Bamman, D. Eisenstein, J. and Schnoebelen, T. (2014) Gender identity and lexical variation in social media. *Journal of Sociolinguistics* 18(2): 135—160.

Barnes, R., Mahar, D., Cockshaw, W. and Wong, I. (2018) Personality and online news commenting behaviours: Uncovering the characteristics of those below the line. *Media International Australia:* 1—14.

Baron, N. S. (2004). See you online: Gender issues in college student use of instant messaging. *Journal of Language and Social Psychology* 23: 397—423.

Bartholomew, D. (1980). Factor Analysis for Categorical Data. *Journal of the Royal Statistical Society. Series B (Methodological) 42*(3): 293—321.

Bastos, M. T., Raimundo, R. L G. and Travitski, R. (2013) Gatekeeping Twitter: Message diffusion in political hashtags. *Media Culture and Society* 35(2): 260—270.

Bawarshi, A. S. & Reiff, M. J. (2010). *Genre: An Introduction to History, Theory, Research and Pedagogy*. Indiana: Parlor Press and The WAC Clearinghouse.

Baym, N. K. (2014) The Perils and Pleasures of Tweeting with Fans. In K. Weller, A. Bruns, J. Burgess, M. Mahrt and C. Pushmann (Eds.) *Twitter and Society*, pp. 221-236. New York: Peter Lang.

BBC News (2014) Troll Peter Nunn guilty of MP Stella Creasy rape tweets. Retrieved on 6 September 2017 from http://www.bbc.co.uk/news/uk-england-london-29034943.

Beers, B. (2019). How Does Twitter Make Money? Retrieved on 8th August 2019 from https://www.investopedia.com/ask/answers/120114/how-does-twitter-twtr-make-money.asp.
Bell, A. (1984) Language style as audience design. *Language in Society* 13(2): 145—204.
Benzécri, J.P. (1979) Sur le calcul des taux d'inertie dans l'analyse d'un questionnaire. *Cahiers de l'Analyse des Données* 4: 377—378.
Benzécri, J.P. (1992) *Correspondence analysis handbook.* New York: Dekker.
Bértoli-Dutra, P. (2014). Multi-Dimensional analysis of pop songs. In T. B. Sardinha & M. V. Pinto (Eds.), *Multi-Dimensional Analysis, 25 Years on: A Tribute to Douglas Biber,* pp. 149—175. Amsterdam/Philadelphia: John Benjamins.
Besnier, N. (1988) The linguistic relationships of spoken and written Nukulaelae registers. *Language* 64: 707—736.
Biber, D. (2006) Stance in spoken and written university registers. *Journal of English for Academic Purposes* 5: 97—116.
Biber, D. (1984) *A model of textual relations within the written and spoken modes.* Ph.D. Dissertation, University of Southern California.
Biber, D. (1985) Investigating macroscopic textual variation through multi-feature/multi-dimensional analyses. *Linguistics* 23(2): 337—60.
Biber, D. (1986) Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language* 62: 384—414.
Biber, D. (1988) *Variation across speech and writing*. Cambridge: Cambridge University Press.
Biber, D. (1989) A typology of English texts. *Linguistics* 27: 3—43.
Biber, D. (1993) Representativeness in corpus design. *Literary and Linguistic Computing* 8(4): 243—257.
Biber, D. (2001) Dimensions of variation among 18th century registers. In H.J. Diller and M. Gorlach (eds.), *Towards a history of English as a history of genres*, pp. 89-110. Heidelberg: C. Winter.
Biber, D. (2004) Modal use across registers and time: an analysis based on the ARCHER corpus. In A. A. Curzan and K. Emmons (Eds.) *Studies in the history of the English language II: Unfolding conversations*, pp. 189—216. Berlin: Mouton de Gruyter.
Biber, D. (2004). Conversation text types: A multi-dimensional analysis. *JADT* 7: 15—34.
Biber, D. (2014) Multi-Dimensional analysis: A personal history. In T. B. Sardinha and M. V. Pinto (eds.) *Multi Dimensional Analysis, 25 years on: A Tribute to Douglas Biber,* pp.xxvi—xxxviii. Amsterdam: John Benjamins Publishing Company.
Biber, D. (2019) Multi-Dimensional Analysis: A Historical Synopsis. In T. B. Sardinha and M. V. Pinto (Eds.) *Multi-Dimensional Analysis: Research Methods and Current Issues*, pp. 11—26. London: Bloomsbury.
Biber, D. and Egbert, J. (2016) Register variation on the searchable web: A multi-dimensional analysis. *Journal of English Linguistics*: 44(2) 95—137.
Biber, D. and Hared, M. (1992) Dimensions of Register Variation in Somali. *Language Variation and Change* 4: 41—75.
Biber, D. and Hared, M. (1994) Linguistic correlates of the transition to literacy in Somali: Language adaptation in six press registers. In D. Biber and

E. Finegan (Eds.), *Sociolinguistic Perspectives on Register*, pp. 182—216. Oxford: Oxford University Press.
Biber, D. and Kurjian, J. (2007). Towards a taxonomy of web registers and text types: A multidimensional analysis. *Language and Computers* 59(1): 109—130.
Biber, D. and Tracy-Ventura, N. (2007) Dimensions of register variation in Spanish. In, G. Parodi (ed.) *Working with Spanish Corpora,* pp. 54-89.
Biber, D. E., & Jones, J. K. (2005). Merging corpus linguistic and discourse analytic research goals: Discourse units in biology research articles. *Corpus Linguistics and Linguistic Theory 1*(2): 151—182.
Biber, D., & Gray, B. (2013). Discourse characteristics of writing and speaking task types on the TOEFL iBT: A lexico-grammatical analysis. *ETS TOEFL Research Report Series*. TOEFL ibT-19.
Biber, D., Davies, M., Jones, J. K., & Tracy-Ventura, N. (2006) Spoken and written register variation in Spanish: A multi-dimensional analysis. *Corpora* 1(1): 1—37.
Biber, D., Egbert, J. and Zhang, M. (2017) Using corpus-based analysis to study register and dialect variation on the searchable web. In E. Friginal (ed.) *Studies in corpus-based sociolinguistics, pp*. 83—111. New York: Routledge.
Biber, D., Johannsson, S. and Leech, G. (1999) *The Longman Grammar of Spoken and Written English.* London: Longman.
Blei, D. M., Ng, A. Y. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3: 993—1022.
Bohmann, A. (2017) Variation in English world-wide: Varieties and genres in a quantitative perspective. *PhD Dissertation*. The University of Texas at Austin
Bouvier, Gwen. 2012. How Facebook Users Select Identity Categories for Self-Presentation. *Journal of Multicultural Discourses* 7 (1): 37—57.
boyd, d. m. & Ellison, N. B. (2007) Social network sites: Definition, history and scholarship. *Journal of Computer-Mediated Communication* 13. Available at: http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html.
boyd, d., Golder, S. & Lotan, G. (2010). Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In *Proceedings of the 43rd Hawaii International Conference on System Sciences* (HICSS'10), pp.1—10. Washington, DC: IEEE Computer Society (January 5-8, 2010).
Brocardo, M. L., Traore, I., Saad, S., and Woungang, I. (2013). Authorship verification for short messages using stylometry. In *Proceedings of the 2013 International Conference on Computer, Information and Telecommunication Systems* (CITS), pp. 1—6. Athens, Greece (May 7-8, 2013).
Brown, P. F., Pietra, V. J. D., deSouza, P. V., Lai, J. C. and Mercer, R. L. (1992) Class-based n-gram models of natural language. *Computational Linguistics* 18(4): 467—479.
Brown, P., & Fraser, C. (1979) Speech as a marker of situation. In K. R. Scherer & H. Giles (Eds.), *Social markers in speech*, pp. 33—62. Cambridge: Cambridge University Press.

Bruns, A., Highfield, T., & Burgess, J. (2013) The Arab Spring and Social Media Audiences: English and Arabic Twitter Users and Their Networks. *American Behavioral Scientist*: 57(7): 871—898.
Bucholtz, M. and Hall, K. (2005) Identity and interaction: a sociocultural linguistic approach. *Discourse Studies* 7(4-5): 585—614.
Buckels, E. E., Trapnell, P. D. and Paulhus, D. L. (2014) Trolls just want to have fun. *Personality and Individual Differences* 67: 97—102.
Buglass, S. L., Binder, J. F., Betts, L. R., & Underwood, J. D. (2016) When ‘friends’ collide: Social heterogeneity and user vulnerability on social network sites. *Computers in Human Behavior* 54: 62—72.
Buller, D. B. and Burgoon, J. K. (1996). Interpersonal Deception Theory. *Communication Theory* 6(3): 203—242.
Bulman, M. (2016) Donald Trump’s ‘celebrity-style’ tweets helped him win US presidential election, says data scientist. Retrieved on 27th August 2019 from https://www.independent.co.uk/news/world/americas/donald-trump-twitter-account-election-victory-president-elect-david-robinson-statistical-analysis-a7443071.html
Burger, J. M., Girgis, Z. M. and Manning, C. C. (2011) In Their Own Words: Explaining Obedience to Authority Through an Examination of Participants’ Comments. *Social Psychological and Personality Science*: 1—7.
Burnap, P. & Williams, M. L. (2014) Hate Speech, Machine Classification and Statistical Modelling of Information Flows on Twitter: Interpretation and Communication for Policy Decision Making. In *Proceedings of the Internet, Policy and Politics Conference* (IPP’14). Oxford, UK. (September 26, 2014).
Burnap, P. & Williams, M. L. (2015) Cyber Hate Speech on Twitter: An Application of Machine Classification and Statistical Modeling for Policy and Decision Making. *Policy & Internet*: 1—20.
Burnap, P. and Williams, M. L. (2016) Us and them: identifying cyber hate on Twitter across multiple protected characteristics. *EPJ Data Science* 5(11): 1—15.
Burnett, G. and Bonnici, L. (2003) Beyond the FAQ: Explicit and implicit norms in Usenet newsgroups. *Library & Information Science Research* 25(3): 333—351.
Butt, D., Fahey, R., Feez, S., Spinks, S. and Yallop, C. (2003) *Using Functional Grammar – An Explorer’s Guide*. Sydney: Macquarie University, National Centre for English Language Teaching and Research.
Carr, N. (2011). *The shallows: What the Internet is doing to our brains*. New York, NY: W.W. Norton & Company.
Carroll, J.B. (1960) Vectors of Prose Style. In T.A. Sebeok (ed.) *Style in Language*, pp. 283—92. Cambridge, MA: MIT Press.
Chafe, W. (1982) Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.) *Spoken and Written Language: Exploring Orality and Literacy*, pp.35—53. Norwood, NJ: Ablex
Chafe, W. (1985) Linguistic differences produced by differences between speaking and writing. In D. R. Olson, N. Torrance, and A. Hildyard (Eds.), *Literature, language, and learning: The nature and conse-*

quences of reading and writing, pp. 105—123. Cambridge: Cambridge University Press.

Chafe, W. and Danielwicz, J. (1987) Properties of Spoken and Written Language. In R. Horowitz and F. J. Samuels (Eds.) *Comprehending Oral and Written Language*, pp. 83—113. New York: Academic Press

Chandrasekharan E., Samory M., Srinivasan A., Gilbert E. (2017) The bag of communities: identifying abusive behavior online with preexisting internet data. In *Proceedings of the 2017 CHI conference on human factors in computing systems (CHI '17, ACM*), pp 3175–3187. New York, NY (May 6-11, 2017).

Chen, M. (2018) A comparison of the situational and linguistic features of high-profile criminal trials and TV series courtroom trials. *Doctoral thesis*, Northern Arizona University.

Chen, Y., Zhu, S., Zhou, Y. & Xu, H. (2012). Detecting Offensive Language in Social Media to Protect Adolescent Online Safety. In *Proceedings of the International Conference on Social Computing* (SocialCom), pp. 71—80. Amsterdam, Netherlands (September 3-5, 2012).

Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C. & Leskovec, J. (2017). Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions. In *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'17).* Portland, Oregon (February 25 – March 1, 2017).

Cheng, J., Danescu-Niculescu-Mizil, C. & Leskovec, J. (2015). Antisocial behavior in online discussion and communities. In *Proceedings of the 9th International AAAI Conference on Web and Social Media* (ICWSM). Oxford, UK (May 26-29, 2015).

Chiluwa, I., & Ifukor, P. (2015). 'War against our Children': Stance and evaluation in #BringBackOurGirls campaign discourse on Twitter and Facebook. *Discourse & Society*, *26*(3): 267—296.

Cho, D., & Kwon, K. H. (2015). The impacts of identity verification and disclosure of social cues on flaming in online user comments. *Computers in Human Behavior 51:* 363—372.

Clarke, I. (2018) Stylistic variation in Twitter trolling. In J. Golbeck (ed.) *Online Harassment: Human-Computer Interaction Series,* pp.151—178. Manhattan, N.Y.: Springer International Publishing.

Clarke, I. (2019) Functional linguistic variation in Twitter trolling. International Journal of Speech Language and the Law 26(1): 1—28.

Clarke, I. and Grieve, J. (2017) Dimensions of abusive language on Twitter. In *Proceedings of the First Workshop on Abusive Language Online,* pp.1--10. Vancouver, Canada (August 4, 2017).

Clarke, I. and Grieve, J. (2019) Stylistic variation on the Donald Trump Twitter account: A linguistic analysis of tweets posted between 2009 and 2018. PlosOne.

Clement, J. (2019) Twitter - Statistics & Facts. Retrieved on 6th August from https://www.statista.com/topics/737/twitter/

Coats, S. (2016) Grammatical feature frequencies of English on Twitter in Finland. In L. Squires (Ed.) *English in computer-mediated communication: Variation, representation, and change,* 179—210. Berlin: de Gruyter Mouton.

Cockerell, J. (2014) Twitter ‘trolls’ Isabella Sorley and John Nimmo jailed for abusing feminist campaigner Caroline Criado-Perez. The Independent Online [Friday 24 January 2014] Retrieved on 6 September 2017 from http://www.independent.co.uk/news/uk/crime/twitter-trolls-isabella-sorley-and-john-nimmo-jailed-for-abusing-feminist-campaigner-caroline-criado-9083829.html

Cocos, A. and Callison-Burch, C. (2017) The Language of Place: Semantic Value from Geospatial Context. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pp. 99—104. Valencia, Spain (April 3-7, 2017).

Cohen, C. (2014) Twitter trolls: The celebrities who've been driven off social media by abuse. Retrieved on 8th August 2019 from https://www.telegraph.co.uk/women/womens-life/11238018/Celebrity-Twitter-trolls-The-famous-people-whove-been-driven-off-social-media-by-abuse.html

Cohen, A. (1999) Generics, Frequency Adverbs, and Probability. *Linguistics and Philosophy* 22: 221—253.

Cohen, R. and Ruths, D. (2013) Classifying political orientation on twitter: It’s not easy! In Proceedings of the Seventh International Association for the Advancement of Artificial Intelligence AAAI Conference on Weblogs and Social Media, pp. 91—99. Cambridge, Massachusetts, USA, (July 8–11, 2013).

Coles, B. A. & West, M. (2016) Weaving the internet together: Imagined communities in newspaper comment threads. *Computers in Human Behavior* 60: 44—53.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014) Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication 64*(2): 317—332.

Collins, P. (2008) The Progressive Aspect in World Englishes: A Corpus-based Study, Australian. *Journal of Linguistics* 28(2): 225—249.

Collot, M., & Belmore, N. (1996) Electronic language: A new variety of English. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*, pp. 13—28. Amsterdam: Benjamins

Communications Act (2003) Retrieved on 8 August 2017 from http://www.legislation.gov.uk/ukpga/2003/21/contents

Condi de Souza, R. (2014) Dimensions of variation in TIME magazine. In T. B. Sardinha and M. V. Pinto (eds.) *Multi Dimensional Analysis, 25 years on: A Tribute to Douglas Biber,* pp.177--194. Amsterdam: John Benjamins Publishing Company

Connor-Linton, J. (1999) Competing Communicative Styles and Crosstalk: A Multi-Feature Analysis. *Language in Society 28*(1): 25—56.

Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F. and Flammini, A. (2011) Political polarization on Twitter. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pp. 89—96. Barcelona, Spain (July 17–21, 2011).

Cook, C., Schaafsma, J. and Antheunis, M. (2018) Under the bridge: An in-depth examination of online trolling in the gaming context. *New Media & Society* 20(9): 3323—3340.
Cook, G. (2001). *The discourse of advertising. 2nd edition.* London, UK: Routledge.
Costa, P. S., Santos, N. C., Cunha, P., Cotter, J. and Sousa, N. (2013) The Use of Multiple Correspondence Analysis to Explore Associations between Categories of Qualitative Variables in Healthy Ageing. *Journal of Aging Research*: 1—14.
Coulthard, M., Johnson, A., and Wright, D. (2017). *An Introduction to Forensic Linguistics*. London: Routledge.
Coyne, I.T., Chesney, B.L. and Madden, N. (2009) Griefing in a virtual community: an exploratory survey of Second Life residents. *Journal of Psychology* 217(4): 214—221.
Craker, N., & March, E. (2016) The dark side of Facebook®: The Dark Tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences 102:* 79—84.
Crossley, S. Allen, L. K. and McNamara, D. S. (2014) A Multi-Dimensional analysis of essay writing. In T. B. Sardinha and M. V. Pinto (eds) *Multi Dimensional Analysis, 25 years on: A Tribute to Douglas Biber* 197—238. Amsterdam: John Benjamins Publishing Company.
Crosthwaite, P. (2016) A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes* 22:166—178.
Cruz, A. G. B., Seo, Y. and Rex, M. (2018) Trolling in online communities: A practice-based theoretical perspective. *The Information Society* 34(1): 15—26.
Cui, Y. and Zhao, Y. (2013) The Use of Second-Person Reference in Advertisement Translation with Reference to Translation between Chinese and English. *International Journal of Society, Culture and Language* 25—36.
Dadvar, M., de Jong, F., Ordelman, R. and Trieschnigg, D. (2012). Improved Cyberbullying Detection Using Gender Information. In Proceedings of the Twelfth Dutch-Belgian Information Retrieval Workshop (DIR 2012), pp. 23—25. Ghent, Belgium (24 February, 2012).
Dadvar, M., Trieschnigg, D. and de Jong. F. (2013) Expert knowledge for automatic detection of bullies in social networks. In *Proceedings of the 25th Benelux Conference on Artificial Intelligence (BNAIC 2013),* pp.57—64. Delft, the Netherlands (7-8 November, 2013).
Dahlberg, L. (2001) Computer-Mediated Communication and the Public Sphere: A Critical Analysis. *Journal of Computer-Mediated Communication* 7(1): 1—19.
Danescu-Niculescu-Mizil, C. Gamon, M. and Dumais, S. (2011) Mark My Words! Lingui stic Style Accommodation in Social Media. In *Proceedings of the International World Wide Web Conference Committee (IW3C2)*, pp. 141—150. Hyderabad, India (March 28–April 2011).
Danescu-Niculescu-Mizil, C. Lee, L. Pang, B. and Kleinberg, J. (2012) Echoes of power: Language effects and power differences in social interaction.

In *Proceedings of the 21st international conference on World Wide Web,* pp. 699–708. Lyon, France (April 16-20, 2012).

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J. and Potts, C. (2013). A computational approach to politeness with application to social factors. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 250—259 Sofia, Bulgaria (August 4-9, 2013).

Darics, E. (2010) Politeness in computer-mediated discourse of a virtual team. *Journal of Politeness Research Language Behaviour Culture* 6(1): 129—150.

Davidson, T., Warmsley, D., Macy, M. Weber, I. (2017) Automated Hate Speech Detection and the Problem of Offensive Language. In *Proceedings of the 11th International AAAI Conference on Web and Social Media* ICWSM'17. Montreal, Canada (May 15-18, 2017).

Dayter, D. (2015) Small stories and extended narratives on Twitter. *Discourse, Context & Media* 10: 19—26.

de Quetteville, H. (2018) How Russian troll factories used Twitter to discredit life-saving vaccines. Retrieved on 8th June 2019 from https://www.telegraph.co.uk/global-health/terror-and-security/inside-story-russian-troll-factories-using-twitter-discredit/

De seta 2013

de Seta, G. (2013). Spraying, fishing, looking for trouble: The Chinese Internet and a critical perspective on the concept of trolling. *The Fibre Culture Journal* 22: 301—317.

Debord, G-E. (1967) *The Society of the Spectacle*, Translated by F. Perlman and J. Supak (Black & Red, 1970; rev. ed. 1977). Retrieved on 12th September 2019 from Library.nothingness.org

Demirhan, K. and Çakir-Demirhan, D. (2014) Gender and politics: Patriarchal discourse on social media. *Public Relations Review* 41(2): 308—310.

Derks, D., Bos, A. E .R., von Grumbkow, J. (2008) Emoticons in Computer-Mediated Communication: Social Motives and Social Context. *CyberPsychology and Behaviour* 11(1):99—101.

Dibbell, J. (2009) Mutilated furries, flying phalluses: put the blame on griefers, the sociopaths of the virtual world. In S. Johnson (Ed.) *The Best Technology Writing 2009*, pp. 9—19. Grand Rapids, MI: University of Michigan Press.

Diener, E. (1980) Deindividuation: The absence of self-awareness and self-regulation in group members. In P. B. Paulus (Ed.) *Psychology of group influence*, pp. 209-242. Hillsdale, NJ: Eribaum.

Dinakar K, Reichart R, Lieberman H (2011) Modeling the detection of textual cyberbullying. In Proceedings of the Fifth International AAAI Conference on Web and Social Media, pp. 11—17. Barcelona, Spain (July 17-21, 2011).

Ditrich, L. & Sassenberg, K. (2017) Kicking out the trolls – Antecedents of social exclusion intentions in Facebook groups. *Computers in Human Behavior* 75: 32—41.

Djuric, N., Zhou, J. Morris, R., Grbovic, M. Radosavljevic, V. and Bhamidipati, N. (2015) Hate speech detection with comment embeddings. In *Proceedings of the 24th International Conference on World Wide Web*, pp. 29–30. Florence, Italy (May 18-22, 2015).

Donath, J. S. (1999) Identity and deception in the virtual community. In P. Kollock and M. A. Smith (eds) *Communities in Cyberspace,* pp. 27—59. London: Routledge

Donoso, G. and Sanchez, D. (2017) Dialectometric analysis of language variation in Twitter. In *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial),* pp. 16—25. Valencia, Spain (April 3, 2017).

Dorsey, J. (2019) How Twitter Needs to change. Retrieved on 1st June 2019 from https://www.ted.com/talks/jack_dorsey_how_twitter_needs_to_change

Dresner, E. and Herring, S.C. (2010) Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication Theory* 20: 249–268.

Dumenco, S. (2011) How Twitter can stop its descent into a cable-news-style disinformation network. *Advertising Age* 82(15): 14—17.

Eckert, P. (1989) *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.

Egbert, J, (2012). Style in nineteenth century fiction: A Multi-Dimensional analysis. *The Scientific Study of Literature* 2(3): 167—198.

Egbert, J. (2019) Corpus Design and Representativeness. In T. B. Sardinha and M. V. Pinto (Eds.) *Multi-Dimensional Analysis: Research Methods and Current Issues*, pp. 27-42. London: Bloomsbury

Eggins, S. (2004) *Introduction to Systemic Functional Linguistics*. 2nd Ed. New York: Continuum.

Ehrhardt, S. (2007). Forensic Linguistics / Authorship Identification In N. N. Daéid & M. M. Houck (Eds.), *Interpol's Forensic Science Review*, pp.363—366. Boca Raton, FL: CRC Press

Eisenstein, J. (2013) What to do about bad language on the internet. In *Proceedings of NAACL-HLT 2013*, pp. 359–369. Atlanta, Georgia (9 – 14 June 2013).

Eisenstein, J. Smith, N. A., and Xing, E. P. (2011) Discovering sociolinguistic associations with structured sparsity. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pp. 1365—1374. Portland, Oregon (June 19-24, 2011).

Eisenstein, J., O'Connor, B., Smith, N. A. and Xing, E. P. (2010) A Latent Variable Model for Geographic Lexical Variation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 1277—1287. MIT, Massachusetts, USA (9-11 October 2010).

Eisenstein, J., O'Connor, B., Smith, N. A. and Xing, E. P. (2014) Diffusion of Lexical Change in Social Media. *Plos One* 9(11): 1—13.

Eleta, I. and Golbeck, J. (2014) Multilingual use of Twitter: Social networks at the language frontier. *Computers in Human Behavior* 41: 424—432.

eMarketer.com (2016) Twitter's User Base to Grow by Double Digits This Year. Retrieved on 31st January 2018 from https://www.emarketer.com/Article/Twitters-User-Base-Grow-by-Double-Digits-This-Year/1014243

Ervin-Tripp, S. (1972) Sociolinguistic Rules: Alteration and Co-Occurrence. In: J. Gumperez, & D. Hymes (Eds.) *Directions in Sociolinguistics,* pp. 213—250. Cambridge: Basil Blackwell.

Fairclough, N. (1992) *Discourse and Social Change*. Blackwell Publishing Ltd., Cambridge.

Festinger, L., Pepitone, A., & Newcomb, T. (1952) Some consequences of de-individuation in a group. *The Journal of Abnormal and Social Psychology 47*:

Fiesler, C., & Proferes, N. (2018) “Participant” Perceptions of Twitter Research Ethics. *Social Media + Society*: 1—14.

Finegan, E. (1980) *Attitudes toward English Usage: The History of a War of Words*. Teachers College Press, Columbia University

Finegan, E. (1982) Form and function in testament language. In R. J. Di Pietro (ed.) *Linguistics and the professions*, pp. 113—120. Norwood, N.J.: Ablex.

Fokin, A. (2016) Trolling and Russia’s Military Strategy. In A. Spruds, A. Rožukalne, K. Sedlenieks, M. Daugulis, D. Potjomkina, B. Tölgyesi, I. Bruge (eds.) *Internet Trolling as a Tool of Hybrid Warfare: The Case of Latvia*, pp. 6—21. NATO stratcom.

Foo, C.Y. and Koivisto, E.M.I. (2004) Defining grief play in MMORPGs: player and developer perceptions. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE ’04),* pp. 245—250. Singapore (June 03-04, 2004).

Fox, J. and Tang, W. Y. (2013) Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior* 33: 314—320.

Freelon, D. and McIlwain, C. D. and Clark, M. (2016) *Beyond the Hashtags: #Ferguson, #Blacklivesmatter, and the Online Struggle for Offline Justice.* Center for Media & Social Impact, American University. Available at SSRN: https://ssrn.com/abstract=2747066 or http://dx.doi.org/10.2139/ssrn.2747066

Friginal, E. (2009). *The Language of Outsourced Call Centers: A Corpus-Based Study of Cross Cultural Interaction*. Amsterdam: John Benjamins Publishing Company.

Friginal, E., Waugh, O. and Titak, A. (2018) Linguistic variation in Facebook and Twitter posts. In E. Friginal (ed.) *Studies in Corpus-Based Sociolinguistics*, pp. 342—363. New York: Routledge.

Gabbatt, A. (2017). Reddit user who created anti-CNN gif used by Trump says sorry for racist posts. Retrieved on 6th August 2019 from: https://www.theguardian.com/us-news/2017/jul/05/reddit-user-hanassholesolo-gif-trump-cnn-tweet

Gaus, A. (2012) Trolling attacks and the need for new approaches to privacy torts. *University of San Francisco Law Review* 47: 353—376.

Gentry, J. (2016) ‘twitteR’ package. Retrieved on 21 November 2017 from https://cran.r-project.org/web/packages/twitteR/twitteR.pdf

Gilbert, E. (2012) Phrases that signal workplace hierarchy. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pp.1037—1046. Seattle, Washington, USA (February 11 - 15, 2012).

Giles, H. (1984) The dynamics of speech accommodation. *International Journal of the Sociology of Language 46.*

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, &

N. Coupland (Eds.), *Studies in emotion and social interaction. Contexts of accommodation: Developments in applied sociolinguistics*, pp. 1—68. New York, NY, US: Cambridge University Press

Gimpel, K., Schneider, N., O'Connor, B., Das, D., Mills, D., Eisenstein, J., Heilman, M., Yogatama, D., Flanigan, J and Smith, N. A. (2011) Part-of-Speech tagging for Twitter: Annotation, features, and experiments. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Short Papers,* pp.19—24. Portland, Oregon (June 19-24, 2011).

Gitari, N. D., Zuping, Z., Damien, H. and Long, J. (2015) A lexicon-based approach for hate speech detection. *International Journal of Multimedia and Ubiquitous Engineering* 10(4): 215–230.

Gligorić, K. Anderson, A. and West, R. (2018) How Constraints Affect Content: The Case of Twitter's Switch from 140 to 280 Characters. In *Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM*), pp. 596—599. Palo Alto, California, USA (June 25-28, 2018).

Glynn, D. (2009). Polysemy, syntax, and variation: A usage-based method for Cognitive Semantics" In V. Evans & S. Pourcel (Eds.) *New directions in Cognitive Linguistics*, pp.77—106. Amsterdam: John Benjamins

Goel, R., Soni, S., Goyal, N., Paparrizos, J., Wallach, H., Diaz, F. and Eisenstein, J. (2016) The Social Dynamics of Language Change in Online Networks. In Proceedings of the International Conference on Social Informatics (SocInfo16), pp. 41—57. Bellevue, WA, USA (November 11-14, 2016)

Golbeck, J., Ashktorab, Z., Banjo, R. O., Berlinger, A., Bhagwan, S., Buntain, C., Cheakalos, P., Geller, A. A., Gergory, Q., Gnanasekaran, R. K., Gunasekaran, R. R., Hoffman, K. M., Hottle, J., Jienjitlert, V., Khare, S., Lau, R., Martindale, M. J., Naik, S., Nixon, H. L., Ramachandran, P., Rogers, K. M., Rogers, L., Sarin, M. S., Shahane, G., Thanki, J., Vengataraman, P., Wan, Z., Wu, D. M. (2017) A Large Human - Labeled Corpus for Online Harassment Research. In Proceedings of WebSci'17, pp. 229—233. Troy, NY, USA (June 25-28, 2017).

Golder, S. A. and Donath, J. (2004) 'Social roles in electronic communities'. Paper presented at the Association of Internet Researchers (AoIR) Conference Internet Research 5.0. Brighton, England (September 19-22, 2004).

Golf-Papez, M. and Veer, E. (2017) Don't feed the trolling: rethinking how online trolling is being defined and combated. *Journal of Marketing Management* 33(15-16): 1336—1354.

Gottfried, J. A., Hardy, B. W., Holbert, R. L., Winneg, K. M. and Jamieson, K. H. (2017) The Changing Nature of Political Debate Consumption: Social Media, Multitasking, and Knowledge Acquisition. *Political Communication* 34(2): 172—199.

Goźdź-Roszkowski, S. (2011) *Patterns of Linguistic Variation in American Legal English: A Corpus-based Study*. Oxford: Peter Lang

Grant, T. (2013) TXT 4n6: method, consistency and distinctiveness in the analysis of SMS text messages. *Journal of Law and Policy* 21(2): 467–494.

Gray, B. (2013) More than Discipline: Uncovering Multi-Dimensional Patterns of Variation in Academic Research Articles". *Corpora* 8(2): 153—181.

Gray, B. (2019) Tagging and Counting Linguistic Features for Multi-Dimensional Analysis. In T. B. Sardinha and M. V. Pinto (Eds.) *Multi-Dimensional Analysis: Research Methods and Current Issues*, pp. 43—66. London: Bloomsbury

Gray, B. E. (2011) Exploring academic writing through corpus linguistics: When discipline tells only part of the story. *PhD Dissertation*. Northern Arizona University.

Gray, B., Biber, D. and Hiltunen, T. (2011) The expression of stance in early (1665-1712) publications of the Philosophical Transactions and other contemporary medical prose. In I. Taavitsainen and P. Pahta (eds.), *Medical Writing in Early Modern English*, pp. 221—247. Cambridge: Cambridge University Press.

Gray, K. L. (2012a) Deviant bodies, stigmatized identities, and racist acts: examining the experiences of African-American gamers in Xbox Live. *New Review of Hypermedia and Multimedia* 18(4): 261—276.

Gray, K. L. (2012b). Diffusion of Innovation Theory and Xbox Live: Examining Minority Gamers' Responses and Rate of Adoption to Changes in Xbox Live. *Bulletin of Science Technology & Society* 32(6): 463—470.

Greenacre, M. & Pardo, R. (2006) Multiple Correspondence Analysis of a Subset of Response Categories. In M. Greenacre and J. Blasius (Eds.), *Multiple Correspondence Analysis and Related Methods*, pp. 197—217. Boca Raton, FL: Chapman & Hall/CRC Press.

Greevy, E. and Smeaton, A. F. (2004) Classifying Racist Texts Using a Support Vector Machine. In SIGIR'04 Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 468—469. Sheffield, United Kingdom (July 25-29, 2004).

Grieve, J. (2014a) A comparison of statistical methods for the aggregation of regional linguistic variation. In B. Szmrecsanyi and B. Wälchli (Eds.) *Aggregating dialectology, typology, and register analysis: Linguistic variation in text and speech*. Berlin/New York: Walter de Gruyter.

Grieve, J**.** (2014b) A multidimensional analysis of regional variation in American English. In T. B. Sardinha and M. V. Pinto (Eds.) *Multi-Dimensional Analysis, 25 years on: A Tribute to Douglas Biber*, pp.3—33. Amsterdam: John Benjamins.

Grieve, J. (2016) *Regional Variation in Written American English.* Cambridge University Press.

Grieve, J., Biber, D., Friginal, E. & Nekrasova, T. (2010). Variation among blogs: A multi dimensional analysis. In A. Mehler, S. Sharoff & M. Santini (Eds.) *Genres on the web: Computational models and Empirical Studies*, pp. 303—322. New York: Springer-Verlag

Grieve, J., Carmody, E., Clarke, I., Gideon, H., Heini, A., Nini, A. & Waibel, E. (2018a). Attributing the *Bixby Letter* using *n*-gram tracing. Digital Scholarship in the Humanities 34(3): 493—512.

Grieve, J., Nini, A. and Guo, D. (2017a) Analysing lexical emergence in American English Online. *English Language and Linguistics* 21: 99—127.

Grieve, J., Nini, A. and Guo, D. (2018b) Mapping lexical innovation on American social media. *Journal of English Linguistics 46(4): 293—319.*

Grieve, J., Ruette, T., Speelman, S. and Geeraerts, D. (2017b) Social functional variation in conversational Dutch. In E. Friginal (editor) *Studies in Corpus-based Sociolinguistics*. Routledge.

Grieve**,** J., Speelman, D., and Geeraerts, D. (2011) A statistical method for the identification and aggregation of regional linguistic variation. *Language Variation and Change* 23: 193—221.

Guy, I. and Shapira, B. (2018) From Royals to Vegans: Characterizing Question Trolling on a Community Questioning Answering Website. In *Proceedings of SIGIR'18*, pp. 834—844. Ann Arbor, MI, USA (July 8-12, 2018).

Halavais, A. (2014) Structure of Twitter: Social and Technical. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, and C. Puschmann (eds.) *Twitter and Society*, pp. 29—42. Oxford: Peter Lang.

Halliday, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.

Halliday, M. A. K. and Hasan, R. (1976) *Cohesion in English.* London: Longman.

Hardaker (2017). "Flaming and trolling" In C. Hoffman and W. Bublitz (eds.) *Pragmatics of social media*, pp.493—522. Berlin: De Gruyter Mouton.

Hardaker, C. (2010) Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research* 6(2): 215—242.

Hardaker, C. (2013) "Uh....not to be nitpicky,,,,,but...the past tense of drag is dragged, not drug." An overview of trolling strategies. *Journal of Language Aggression and Conflict* 1(1): 58—86.

Hardaker, C. (2015) "I refuse to respond to this obvious troll": An overview of responses to (perceived) trolling. *Corpora* 10(2): 201—229.

Hardaker, C. & McGlashan, M. (2016) "Real men don't hate women": Twitter rape threats and group identity. *Journal of Pragmatics* 91: 80—93.

Hardy, J. A. and Friginal, E. (2012) Filipino and American online communication and linguistic variation. *World Englishes* 31(2): 143—161.

Harper, D. (2001-2017) "troll". Online Etymology Dictionary. Retrieved on 6 September 2017 from http://www.etymonline.com/index.php?term=troll

Hayes, A. (2019) Dotcom Bubble. Retrieved on 8th August 2019 from https://www.investopedia.com/terms/d/dotcom-bubble.asp

Hellmore, E. (2018) Tyler Barriss, accused of making hoax call regrets death of 'swatting' victim. Retrieved on 5th January 2019 from https://www.theguardian.com/us-news/2018/jan/15/tyler-barriss-swatting-death-regret

Herring, S. C. (2002) Computer-mediated communication on the Internet. *Annual Review of Information Science and Technology* 36: 109—168.

Herring, S. C. (2012) Grammar and electronic communication. In C. Chapelle (ed.) *Encyclopedia of applied linguistics.* Hoboken, NJ: Wiley-Blackwell.

Herring, S. C. and Paolillo, J. C. (2006) Gender and genre variation in weblogs. *Journal of Sociolinguistics* 10(4): 439—459.

Herring, S. C., Job-Sluder, K., Scheckler, R. & Barab, S. (2002). Searching for Safety Online: Managing “Trolling” In a Feminist Forum. *The Information Society* 18: 371—384.

Hicks, D. and Gasca, D. (2019) A healthier Twitter: Progress and more to do. Retrieved on 8th August from https://blog.twitter.com/en_us/topics/company/2019/health-update.html

Higgin, T. (2013). “/b/lack up: What Trolls Can Teach Us About Race” *The Fibreculture Journal* 22: 132—150.

Hinduja, S. and Patchin, J. W. (2014) *Cyberbullying: Identification, Prevention and Response*. Cyberbullying Research Center.

Hine, G. E., Onaolapo, J. Cristofaro, E. D., Kourtellis, N. Leontiadis, I., Samaras, R., Stringhini, G. & Blackburn, J. (2017). Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan’s Politically Incorrect Forum and Its Effects on the Web. In *Proceedings of the 11th Inter International AAAI Conference on Web and Social Media* ICWSM’17. Montreal, Canada (May 15-18, 2017).

Hmielowski, J. D., Hutchens, M. J. and Cicchirillo, V. J. (2014) Living in an age of online incivility: examining the conditional indirect effects of online discussion on political flaming. *Information, Communication & Society* 17(10): 1196—1211.

Hogan, B. (2013). Pseudonyms and the Rise of the Real-Name Web In J. Hartley, J. Burgess & A. Bruns (Eds.). *A Companion to New Media Dynamics*, pp. 290—308. Chichester: Blackwell Publishing Ltd.

Holler, J., Kendrick, K. H., Casillas, M. and Levinson, S. C. (2015) Editorial: Turn-Taking in Human Communicative Interaction. F*ront. Psychol*. 6:1919.

Holt, L. (2012) Using laugh responses to defuse complaints. *Research on Language and Social Interaction* 45(4): 430—448.

Honeycutt, C. & Herring, S. C. (2009). Beyond Microblogging: Conversation and Collaboration via Twitter. In *Proceedings of the Forty-Second Hawai’i International Conference on System Sciences* (HICSS-42). Los Alamitos, CA: IEEE Press (5-8 January 2009)

Hosseinmardi, H., Mattson, S. A., Rafiq, R. I., Han, R., Lv, Q. & Mishr, S. (2015) Analyzing Labeled Cyberbullying Incidents on the Instagram Social Network. In T. Y. Liu, C. Scollon & W. Zhu (Eds.). *Social Informatics. Lecture Notes in Computer Science*, pp. 49—66. Switzerland: Springer International Publishing.

Hours, H., Fleury, E., Karsai, M. (2016) Link prediction in the Twitter mention network: impacts of local structure and similarity of interest. In *Proceedings of the 16th IEEE International Conference on Data Mining (ICDM) - DMHAA Workshop*. Barcelona, Spain. (Dec 2016).

Hovy, D. (2015) Demographic Factors Improve Classification Performance. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, pp. 752—762. Beijing, China (July 26-31, 2015).

https://dev.twitter.com/oauth/overview/application-owner-access-tokens

Huang, B., & Raisi, E. (2018). Weak Supervision and Machine Learning for Online Harassment Detection. In J. Golbeck (ed.) *Online Harassment:*

Human-Computer Interaction Series, pp.5—28. Manhattan, N.Y.: Springer International Publishing.

Huang, Y. Guo, D., Kasakoff, A., and **Grieve, J**. (2016) Understanding US regional linguistic variation with Twitter data analysis. *Computers, Environment and Urban Systems* 59: 244—255.

Husson, F., Josse, J. Le, S. and Mazet, J. (2017) Package 'FactoMineR'. Retrieved on 30 June 2017 from http://factominer.free.fr

Hymes, D. (1974) *Foundations in Sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press

Jalayer, M. and Zhou, H. (2017) A multiple correspondence analysis of at-fault motorcycle-involved crashes in Alabama. *Journal of Advanced Transportation* 50(8): 2089—2099.

Jang, S. C. (1998) Dimensions of Spoken and Written Taiwanese: A Corpus-Based Register Study. *PhD Dissertation*. University of Hawaii.

Jappe, A. (1999) *Guy Debord*. University of California Press: Berkley.

Java, A., Song, X., Finin, T. & Tseng, B. (2007). Why We Twitter: Understanding Microblogging Usage and Communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD workshop on Web mining and social network analysis*, pp.56—65. San Jose, CA: ACM. (August 12-15, 2017).

Jefferson, G. (2002) Is 'no' an acknowledgement token? Comparing American and British uses of (+)/(-) tokens. *Journal of Pragmatics* 34(10/11): 1345—1383.

Jones, S. (2013). 'Don't be Rude on the Road': Cycle Blogging, Trolling and Lifestyle. *The Fibre Culture Journal* 22: 233—256.

Jonsson, E. (2015) *Conversational writing: A multidimensional study of synchronous and supersynchronous computer-mediated communication*. Frankfurt am Main: Peter Lang.

Kahma, N. and Toikka, A. (2012) Cultural map of Finland 2007: analysing cultural differences using multiple correspondence analysis. *Cultural Trends* 21(2): 113—131.

Kanoksilapatham, B. (2007) Rhetorical moves in biochemistry research articles. In D. Biber, U. Connor and T. A. Upton (Eds.) *Discourse on the move: Using corpus analysis to describe discourse structure*, pp. 73—119. Amsterdam: John Benjamins Publishing Company

Kaplan, A. and Haenlein, M. (2010) Users of the world, unite! The challenges and opportunities of social media. Business Horizons 53(1): 59--68

Karppi, T. (2013). FCJ-166 'Change name to No One. Like people's status' Facebook Trolling and Managing Online Persona's. *The Fibreculture Journal* 22: 277—299.

Katinskaya, A., & Sharoff, S. (2015). *Applying Multi-Dimensional Analysis to a Russian Webcorpus: Searching for Evidence of Genres*. *BSNLP@RANLP*.

Kearney, M. W. (2018) Package rtweet. Retrieved on 22nd June 2018 from: https://cran.r-project.org/web/packages/rtweet/rtweet.pdf

Kim, Y. and Biber, D. (1994) A corpus-based analysis of register variation in Korean. In D. Biber and E. Finegan (Eds.) *Sociolinguistic Perspectives on Register*: 157—181.

Klempka, A. & Stimson, A. (2014) Anonymous Communication on the Internet and Trolling. *Concordia Journal of Communication Research:* 1—14.

Koban, K., Stein, J., Eckhardt, V., & Ohler, P. (2018). Quid pro quo in Web 2.0. Connecting personality traits and Facebook usage intensity to uncivil commenting intentions in public online discussions. *Computers in Human Behavior 79*: 9—18.

Kopecký, K. (2016). Misuse of web cameras to manipulate children within the so-called webcam trolling. *Telematics and Informatics* 33: 1—7.

Koppel, M., Argamon, S. and Shimoni, A. R. (2002) Automatically Categorizing Written Texts by Author Gender. *Literary and Linguistic Computing* 17(4) 401–412.

Krasodomski-Jones, A. (2016) *Talking to ourselves? Political debate online and the echo chamber effect*. London: Demos

Kress, G. (2010) *Multimodality: A social semiotic approach to contemporary communication*. London: Routledge.

Krishnamurthy, B., Gill, P. & Arlitt, M. (2008). A Few Chirps About Twitter. In *Proceedings of the First Workshop on Online Social Networks*, pp. 19—24. Seattle, WA

Kruger, H. and van Rooy, B. (2016) A multidimensional analysis of translated English and a non-native indigenised variety of English. *English World-Wide* 37(1): 26—57.

Kwak, H., Blackburn, J. and Han, S. (2015) Exploring cyberbullying and other toxic behavior in team competition online games In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 3739—3748. Seoul, Republic of Korea (April 18-23, 2015).

Labov, W. and Waletzky, J. (1967) Narrative analysis. In J. Helm (ed.), *Essays on the Verbal and Visual Arts,* pp. 12—44. Seattle: University of Washington Press.

Lamb, W. (2008) *Scottish Gaelic speech and writing: Register variation in an endangered language*. Belfast: Cló Ollscoil na Banríona

Lamba, H. Malik, M. M. and Pfeffer, J. (2015) A Tempest in a Teacup? Analyzing Firestorms on Twitter. In *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2015)* pp. 17—24. Paris, France (August 25-28, 2015).

Laville,S. (2017) Internet troll who sent Labour MP anti-Semitic messages is jailed. Retrieved on 1 October 2018 from https://www.theguardian.com/uk-news/2017/feb/10/internet-troll-who-sent-labour-mp-antisemitic-messages-is-jailed

Layton, R., Watters, P. and Dazeley, R. (2010) Authorship attribution for twitter in 140 characters or less. In *Proceedings of the 2010 Second Cybercrime and Trustworthy Computing Workshop*, CTC '10, pp 1–8. Washington, DC, USA.

Le Roux, B. & Rouanet, H. (2010). *Multiple Correspondence Analysis*. California: SAGE Publications, Inc.

Le Roux, B. and Rouanet, H. (1984) L'analyse multidimensionnelle des données structurées. *Mathématiques Informatique et Sciences humaines*: 85—118.

Le Roux, B. and Rouanet, H. (1988) Interpreting axes in multiple correspondence analysis: method of the contributions of points and

deviations. In J. Blasius and M. Greenacre (eds.) *Visualization of Categorical Data*, pp. 197—220. Academic Press.
Le Roux, B., Rouanet, H., Savage, M., & Warde, A. (2008). Class and Cultural Division in the UK. *Sociology 42*(6): 1049.
Le, Q. and Mikolov, T. (2014) Distributed Representations of Sentences and Documents. In *Proceedings of the* 31st *International Conference on Machine Learning*. Beijing, China.
Lea, M. and Spears, R. (1991) Computer-mediated communication, deindividuation and group decision-making. *International Journal of Man-Machine Studies* 34: 283—301.
Lea, M., O'Shea, T., Fung, P. and Spears, R. (1992) 'Flaming' in computer-mediated communication. Observations, explanations, implications. In M. Lea (ed.), *Contexts of Computer-Mediated Communication*, pp.89–112. New York: Harvester Wheatsheaf.
Lee, C. K. M. (2011) Micro-blogging and status updates on Facebook: Texts and Practices. In Thurlow, C and Mroczek, K. (eds.) *Digital discourse: Language in the new media*, pp. 111—128. Oxford University Press.
Lee, D. Y. W. (2001). Genres, Registers, Text Types, Domains, and Styles: Clarifying the Concepts and Navigating a Path Through the BNC Jungle. *Language Learning and Technology* 5(3): 37—72.
Lee, Y. Yoon, S. and Jung, K. (2018) Comparative Studies of Detecting Abusive Language on Twitter. In *Proceedings of the 2nd Workshop on Abusive Language Online, ALW@EMNLP 2018*, pp. 101—106. Brussels, Belgium (October 31, 2018).
Leppänen, S., Kytölä, S., Jousmäki, H., Peuronen, S. and Westinen, E. (2014) Entextualization and resemiotization as resources for identification in social media. In P. Seargeant and C. Tagg (eds.) *The language of social media: Communication and community on the internet*, pp. 112—138. Basingstoke: Palgrave.
Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology 68*(1): 151—158.
Lewandowsky, S., Ecker, U. K. H., and Cook, J. (2017) Beyond misinformation: understanding and coping with the post-truth era. *Journal of Applied Research in Memory and Cognition* 6(4): 353—369.
Liao, L., Jiang, J., Ding, Y., Huang, H. and Lim, E-P. (2014) Lifetime lexical variation in social media. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pp. 1643—1649. Québec City, Québec, Canada.
Lim, J. S., Hwang, Y., Kim, S., & Biocca, F. A. (2015). How social media engagement leads to sports channel loyalty: Mediating roles of social presence and channel commitment. *Computers in Human Behavior 46*: 158—167.
Litt, E. (2012) *Knock, Knock*. Who's There? The Imagined Audience. *Journal of Broadcasting & Electronic Media* 56(3): 330—345.
Liu, F., Weng, F. and Jiang, X. (2012) A Broad-Coverage Normalization System for Social Media Language. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pp. 1035–1044. Jeju, Republic of Korea (8-14 July 2012).

Ljosland, R. (2013) “Don’t mention the tail!: A look into the hidden world of the Huldrefolk”. The Orcadian, pp.15 [Thursday 28th November, 2013]. Retrieved on 6 September 2017 from https://www.uhi.ac.uk/en/research-enterprise/cultural/centre-for-nordic-studies/mimirs-well-articles/don2019t-mention-the-tail-a-look-into-the-hidden-world-of-the-huldrefolk/

Lumsden, K. & Morgan, H. M. (2012). ‘Fraping’, ‘Sexting’, ‘Trolling’ and ‘Rinsing’: social networking, feminist thought and the construction of young women as victims or villains. In *Proceedings of Forthcoming Feminisms: Gender Activism, Politics and Theory* (BSA Gender Study Group Conference). Leeds, UK. (26 October, 2017).

Lumsden, K. and Morgan, H. M. (2017) Cyber-Trolling as Symbolic Violence: Deconstructing Gendered Abuse Online. In N. Lombard (ed). *The Routledge Handbook of Gender and Violence,* Chapter 9. London: Routledge.

Lundberg G. J. W., Neel R., Lassetter B. and Todd A. R. (2018) Racial bias in implicit danger associations generalizes to older male targets. *PLoS ONE* 13(6): e0197398.

MacArthur, A. (2019). The Real History of Twitter, in Brief. Retrieved on 6th August 2019 from: https://www.lifewire.com/history-of-twitter-3288854

MacKinnon, R. & Zuckerman, E. (2012) Don’t feed the trolls. *Index on Censorship* 41(4): 14—24.

Mahase, E. (2019) Measles cases rise 300% globally in first few months of 2019. *BMJ* 365

Maity, S. K., Chakraborty, A., Goyal, P. and Mukherjee, A. (2018) Opinion Conflicts: An Effective Route to Detect Incivility in Twitter. In *Proceedings of the ACM on Human-Computer Interaction* 2(117): 1—27.

Maltby, J., Day, L. and Hatcher, R. M. (2015) Implicit theories of online trolling: evidence that attention-seeking conceptions are associated with increased psychological resilience. *British Journal of Psychology* 107(3): 448—466.

Manivannan, V. (2013). Tits or GTFO: The logics of misogyny on 4chan’s Random - /b/. *The Fibreculture Journal* 22: 108—131.

March, E. (2019) Psychopathy, sadism, empathy, and the motivation to cause harm: New evidence confirms malevolent nature of the Internet Troll. *Personality and Individual Differences* 141: 133—137.

March, E., Grieve, R., Marrington, J. & Jonason, P. K. (2017) Trolling on Tinder® (and other dating apps): Examining the Dark Tetrad and impulsivity. *Personality and Individual Differences* 110: 139—143.

Martin, J. R. and White, P. R. R. (2005) *The Language of Evaluation: Appraisal in English*. Basingstoke, Hampshire: Palgrave MacMillan.

Marwick, A. E. and Boyd, D. (2011) The Drama! Teen Conflict, Gossip, and Bullying in Networked Publics (September 12, 2011). A Decade in Internet Time: *Symposium on the Dynamics of the Internet and Society*, September 2011. Available at SSRN: https://ssrn.com/abstract=1926349

McCosker, A. (2013). Productive Provocations: Vitriolic Media, Spaces of Protest and Agonistic Outrage in the 2011 England Riots. *The Fibreculture Journal* 22: 170—192.

McCosker, A. (2014) Trolling as provocation: YouTube's agonistic publics. *Convergence*, *20*(2): 201—217.

McEnery, T., Xiao, R. & Tono, Y. (2006) *Corpus-based Language Studies: An Advanced Resource Book*. London/New York: Routledge.

McGuire, W. J. and McGuire, C. V. (1986) Differences in Conceptualizing Self Versus Conceptualizing Other People as Manifested in contrasting Verb Types Used in Natural Speech. *Journal of Personality and Social Psychology* 51(6): 1135—1143.

McIntosh, S. & Pavlik, J. V. (2011) *Converging Media: A New Introduction to Mass Communication*. Oxford: Oxford University Press

McLaughlin, M. L., Osborne, K. K., & Smith, C. B. (1995). Standards of conduct on Usenet. In S. Jones (Ed.), *Cybersociety: Computer-mediated comnaunication and community,* pp. 90-111. Thousand Oaks, CA: Sage.

Mehdad, Y. & Tetreault, J. (2016) Do Characters Abuse More Than Words? In *Proceedings of the SIGDIAL 2016 Conference*, pp. 299—303. Los Angeles, USA (13-15 September, 2016).

Merritt, E. (2012). An Analysis of the Discourse of Internet Trolling: A Case Study of Reddit.com. *PhD dissertation*. Mount Holyoke College.

Messias, E., Kindrick, K. and Castro, J. (2014) School bullying, cyberbullying, or both: correlates of teen suicidality in the 2011 CDC Youth Risk Behavior Survey. *Comprehensive Psychiatry* 55(5): 1063—1068.

Meyrowitz, J. (1994) Medium theory. *Communication theory today*: 50—77.

Mihaylov, T. & Nakov, P. (2016). Hunting for Troll Comments in News Community Forums. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 399—405. Berlin, Germany (August 7-12, 2016).

Mihaylov, T., Georgiev, G. D. & Nakov, P. (2015) Finding Opinion Manipulation Trolls in News Community Forums. In *Proceedings of the 19th Conference on Computational Language Learning*, pp. 310—314. Beijing, China: ACL (July 30-31, 2015).

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013) Efficient estimation of word representations in vector space. Proceedings of Workshop at ICLR. arXiv:1301.3781v1

Miller, V. (2012) A Crisis of Presence: On-line Culture and Being in the World. *Space and Polity* 16(3): 265—285.

Milner, R. M. (2013) Hacking the Social: Internet Memes, Identity Antagonism, and the Logic of Lulz. *The Fibreculture Journal* 22:61—91.

Milroy, J. and Milroy, L. (1985) Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21: 339—384.

Milroy, L. and Gordon, M. (2003) *Sociolinguistics: Method and Interpretation*. Hoboken, NJ: Wiley-Blackwell.

Mishra, P. Yannakoudakis, H. and Shutova, E. (2018) Neural Character-based Composition Models for Abuse Detection. In *Proceedings of the EMNLP Workshop on Abusive Language Online* 2018, pp. 1—11.

Mislevy, R. J. (1986). Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics* 11(1): 3—31.

Mocanu, D., Rossi, L., Zhang, Q., Karsai, M. and Quattrociocchi, W. (2014) Collective attention in the age of (mis)information. *Computers in Human Behavior*: 1198—1204.
Mojica, L. G. (2017). A Trolling Hierarchy in Social Media and A Conditional Random Field For Trolling Detection https://arxiv.org/pdf/1704.02385.pdf
Morrissey, B. and Yell, S. (2016) Performative trolling: Szubanski, Gillard, Dawson and the nature of the utterance. *Persona Studies* 2(1): 27—40.
Morstatter, F., Pfeffer, J., Liu, H. and Carley, K. M. (2013) Is the sample good enough? Comparing data from Twitter’s streaming API with Twitter’s Firehose. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, pp.400—408.
Mortensen, T. E. (2018). Anger, Fear, and Games: The Long Event of #GamerGate. *Games and Culture* 13(8): 787—806.
Mukherjee, A. and Liu, B. (2010) Improving gender classification of blog authors. In Proceedings of EMNLP '10 Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, pp. 207—217. Cambridge, Massachusetts (October 9-11, 2010).
Munger, K. (2016). Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment. *Political Behavior*: 1—21.
Munger, K. (2017). Don’t @ Me: Experimentally Reducing Partisan Incivility on Twitter. Retrieved on 31st August 2018 from cess.nyu.edu/wp-content/uploads/2017/02/Dont-@-Me.pdf
Murthy, D. and Longwell, S. A. (2013) TWITTER AND DISASTERS. *Information, Communication & Society* 16(6): 837—855.
Neeley-Cohen, M. (2016) NBA Twitter is changing the way we watch sports. Retrieved on 1st February 2019 from: https://newrepublic.com/article/127237/nba-twitter-changing-way-watch-sports
Nguyen, D., & Eisenstein, J. (2017). A Kernel Independence Test for Geographical Language Variation. *Computational Linguistics*, 1—40.
Nguyen, D., Gravel, R., Trieschnigg, D. and Meder, T. (2013) “How old do you think I am?" A study of language and age in Twitter. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, pp. 439—448. Boston, MA.
Nicol, S. (2012) Cyber-bullying and trolling. *Youth Studies Australia* 31(4): 3—4.
Nicotra, J. (2016) *Disgust, distributed: Virtual public shaming as epideictic assemblage. Enculturation*. Retrieved on 1 October 2018 from http://enculturation.net/disgust-distributed
Nini, A. (2017) Register variation in malicious forensic texts. *International Journal of Speech, Language and the Law* 24(1): 99—126.
Nini, A. (2018) An authorship analysis of the Jack the Ripper letters. *Digital Scholarship in the Humanities* 33(3): 621—638.
Nini, A. (2019) The Multi-Dimensional Analysis Tagger. In T. B. Sardinha and M. V. Pinto (Eds.) *Multi-Dimensional Analysis: Research Methods and Current Issues*, pp. 67—94. London: Bloomsbury
Nini, A., Corradini, C., Guo, D. and Grieve, J. (2017) The application of growth curve modeling for the analysis of diachronic corpora. *Language Dynamics and Change* 7(1): 102—125.

Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y. and Chang, Y. (2016) Abusive language detection in online user content. In *Proceedings of the 25th International Conference on World Wide Web*, pp. 145–153. Montréal, Québec, Canada (April 11–15, 2016).

Norrick, N. R. (1994) Involvement and joking in conversation. Journal of Pragmatics 22: 409—430.

Nyst, C. and Monaco, N. Woolley, S. C. (2018) State-sponsored trolling. *Institute for the Future* Retrieved on 8th August 2019 from http://www.iftf.org/fileadmin/user_upload/images/DigIntel/IFTF_State_sponsored_trolling_report.pdf

O'Reilly, T. (2005). *What Is Web 2.0 Design Patterns and Business Models for the Next Generation of Software*. Retrieved on 1st January 2019 from http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

O'Sullivan, P. B. & Flanagin, A. J. (2003). Reconceptualizing flaming and other problematic messages. *New Media and Society* 5(1): 69—94.

Ochs, E. (1979) Planned and unplanned discourse. In T. Givon (ed.) *Syntax and semantics, Vol. 12: Discourse and semantics*. New York: Academic Press

Omernick, E. & Sood, S. O. (2013) The impact of anonymity in online communities. In *Proceedings of the International Conference on Social Computing*, pp. 526—535. Washington D.C. (September 8-14, 2013).

Oremus, W. (2017) Twitter's New Order. Retrieved on 6th July 2019 from: http://www.slate.com/articles/technology/cover_story/2017/03/twitter_s_timeline_algorithm_and_its_effect_on_us_explained.html

Orosz, G., Vallerand, R. J., Bőthe, B., Tóth-Király, I., & Paskuj, B. (2016) On the correlates of passion for screen-based behaviors: The case of impulsivity and the problematic and non-problematic Facebook use and TV series watching. *Personality and Individual Differences* 101: 167–176.

Ott, B. L. (2017). The age of Twitter: Donald J. Trump and the politics of debasement. *Critical Studies in Media Communication* 34(1): 59—68.

Ouvrein, G., Pabian, S., Machimbarrena, J. M., Charlotte J.S. De Backer, C. J. S. and Vandebosch, H. (2018) Online Celebrity Bashing: Wrecking Ball or Good for You? Adolescent Girls' Attitudes Toward the Media and Public Bashing of Miley Cyrus and Selena Gomez. *Communication Research Reports* 35(3): 261—271.

Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., Schneider, N. and Smith, N. A. (2013) Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters. *Proceedings of NAACL-HLT 2013*, pp. 380–390. Atlanta, Georgia (9–14 June 2013).

Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., Schneider, N. and Smith, N. A. (2012) Part-of-speech tagging for Twitter: Word clusters and other advances. Technical Report CMU-ML-12-107, Carnegie Mellon University.

Page, R. (2012) The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse & Communication* 6(2): 181—201.

Parodi, G. (2007) Variation across registers in Spanish: Exploring the El Grial PUCV Corpus'. In G. Parodi (ed.) *Working with Spanish Corpora*, pp. 11—53. London: Continuum.

Passonneau, R. J., Ide, N., Su, S., & Stuart, J. (2014). “Biber Redux: Reconsidering Dimensions of Variation in American English” In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pp. 565—576. Dublin, Ireland (August 23-29, 2014)

Pavalanathan, U. and Eisenstein, J. (2015) Audience-modulated variation in online social media. *American Speech* 90(2): 187—213.

Pavalanathan, U., Fitzpatrick, J., Kiesling, S. F., Eisenstein, J. (2017). A Multidimensional Lexicon for Interpersonal Stancetaking. In *Proceedings of the Association for Computational Linguistics*. Vancouver, Canada: ACL. (July 30 – August 4, 2017).

Pavalanathan, U., Han, X., and Eisenstein, J. (2018) Mind Your POV: Convergence of Articles and Editors Towards Wikipedia’s Neutrality Norm. In Proceedings of the ACM on Human-Computer Interaction 2(137)

Pavlopoulos, J., Malakasiotis, P. and Androutsopoulos, I. (2017) Deeper Attention to Abusive User Content Moderation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp.1125—1135. Copenhagen, Denmark (September 7–11, 2017).

Petitjean, C. & Morel, E. (2017) “Hahaha”: Laughter as a resource to manage WhatsApp conversations 111: 1—19.

Petyko, M. (2017). *“You’re trolling because…” – A Corpus-based Study of Perceived Trolling and Motive Attribution in the Comment Threads of Three British Political Blogs*. 56-60. https://doi.org/10.5281/zenodo.1040875

Pew Research Center (2019) Social Media Fact Sheet. Retrieved on 13th September 2019 from https://www.pewinternet.org/fact-sheet/social-media/

Phillips, W. (2011). Meet the trolls. *Index on Censorship* 40(2): 68—76.

Phillips, W. (2013). The House That Fox Built: Anonymous, Spectacle, and Cycles of Amplification. *Television & New Media* 14(6): 494—509.

Phillips, W. (2016) *This is why we can’t have nice things: Mapping the Relationship between Online Trolling and Mainstream Culture*. Cambridge, MA: The MIT Press

Phillips, W. and Milner, R. M. (2017) The ambivalent internet: mischief, oddity and antagonism online. Malden, MA: Polity Press

Pilkington, E. (2013). Justine Sacco, PR executive fired over racist tweet, ‘ashamed’. The Guardian.com. Retrieved on 7th June 2017 from https://www.theguardian.com/world/2013/dec/22/pr-exec-fired-racist-tweet-aids-africa-apology>

Pinto, M. V. (2014) Dimensions of variation in North American. In T. B. Sardinha and M. V. Pinto (eds) *Multi Dimensional Analysis, 25 years on: A Tribute to Douglas Biber,* pp.109—148. Amsterdam: John Benjamins Publishing Company

Poe, N. (2005) Poe’s Law. Retrieved on 8th August 2019 from https://web.archive.org/web/20170114124412/http://www.christianforums.com/threads/big-contradictions-in-the-evolution-theory.1962980/page-3#post-17606580

Postmes, T., Spears, R., and Lea, M. (2000) The formation of group norms in computer-mediated communication. *Human Communication Research* 26(3): 341—371.

Prabhakaran, V., Rambow, O. and Diab, M. (2012) Predicting Overt Display of Power in Written Dialogs. In *Human Language Technologies: The 2012 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Montreal, Canada. (June, 2012)

Prideaux, S. (2019) New Miss Algeria 2019 targeted by racist trolls. Retrieved on 5th May 2019 from https://www.thenational.ae/lifestyle/new-miss-algeria-2019-targeted-by-racist-trolls-1.810279

Procter, R. Crump, J., Karstedt, S., Voss, A. and Cantijoch, M. (2013) Reading the riots: what *were* the police doing on Twitter? *Policing and Society* 23(4): 413—436.

Quaglio, P. (2009) *Television Dialogue: The Sitcom Friends Vs. Natural Conversation.* Amsterdam: John Benjamins Publishing.

Quirk, R. and Greenbaum, S. and Leech, G. and Svartvik, J. (1985) *A Comprehensive Grammar of the English Language*. London: Longman.

Radulova, L. (2014) Putting your iPhone in the microwave will COOK it, not charge it: Apple urges users not to fall for convincing internet hoax... too late for many https://www.dailymail.co.uk/news/article-2768976/Emergency-services-forced-step-iPhone-users-fall-internet-prank-explains-use-microwave-charge-phone.html

Rahimi, A., Cohn, T., and Baldwin, T. (2017) A neural model for user geolocation and lexical dialectology. In *Proceedings of the Annual Meeting of the Association of Computational Linguistics 2: Short Papers*, pp. 209-216. Vancouver, Canada.

Raisi, E. and Huang, B. (2017) Cyberbullying Detection with Weakly Supervised Machine Learning. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 409—416. Sydney, Australia (July 31- August 3, 2017).

Rao, D., Yarowsky, D., Shreevats, A. and Gupta, M. (2010) Classifying latent user attributes in Twitter. In Proceedings of the *2nd International Workshop on Search and Mining User-Generated Content*. ACM. Toronto, Ontario, Canada.

Reppen, R. (1994). *Variation in elementary student language: A multi-dimensional perspective. Doctoral dissertation,* Northern Arizona University.

Reppen, R. (2001) Register variation in student and adult speech and writing. In S. Conrad, D. Biber (Eds.) *Variation in English: Multi-dimensional studies*, pp. 187—199. Harlow, UK: Longman.

Reppen, R. (2007) First Language & Second Language Writing Development of Elementary Students – Two Perspective – In Y. Kawaguchi, T. Takagaki, N. Tomimori and Y. Tsuruga (eds.) *Corpus-Based Perspectives in Linguistics*, pp.147—168. Amsterdam: John Benjamins Publishing

Rickford, J. R. and Eckert, P. (2001) Introduction. In P. Eckert and J. R. Rickford (eds.) *Style and sociolinguistic variation*, pp.1—18. Cambridge: Cambridge University Press.

Rosenthal, S. and McKeown, K. (2011) Age Prediction in Blogs: A Study of Style, Content, and Online Behavior in Pre- and Post-Social Media

Generations. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pp. 763–772. Portland, Oregon (June 19-24, 2011).

Rösner, L., Winter, S., & Krämer, N.C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior* 58: 461—470.

Rubin, V. L., Camm, S. C. (2011) Griefing and Deception in Video Games: Examining Attitudes towards the Phenomena. In Proceedings of the 39th Annual Conference of the Canadian Association for Information Science: Exploring interactions of people, places and information. Fredericton, N.B., Canada (June 2-4, 2011).

Salo, J. (2016) Teen who killed herself in front of family is still being bullied online. Retrieved on 24th January 2019 from: https://nypost.com/2016/12/13/teen-who-killed-herself-in-front-of-family-is-still-being-bullied-online/

Samghabadi, N. S., Maharjan, S., Sprague, A., Diaz-Sprague, R. and Solorio, T. (2017) Detecting Nastiness in Social Media. In *Proceedings of the First Workshop on Abusive Language Online*, pp. 63—72. Vancouver, Canada (July 30 - August 4, 2017).

Sanfilippo, M. R., Yang, S., and Fichman, P. (2017) Managing online trolling: From deviant to social and political trolls. In *Proceedings of the 50th Hawaii International Conference on System Sciences.* Waikoloa, HI (January 4 - 7, 2017)

Sankaranarayanan, J., Samet, H., Teitler, B. E., Lieberman, M. D., and Sperling, J. (2009) Twitterstand: news in tweets. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '09,* pp. 42—51. New York, NY, USA

Santini, M. (2005). Automatic Text Analysis: Gradations of Text Types in Web Pages. In *Proceedings of the 10th ESSLLI Student Session,* pp. 276-285. Baden-Württemberg, Germany.

Sardinha, T. B. (2014) Comparing Internet and pre-Internet registers. In T. B. Sardinha and M. V. Pinto (eds.) *Multi-Dimensional Analysis, 25 years on: A Tribute to Douglas Biber*, pp. 81—108. Amsterdam/Philiadelphia: John Benjamins.

Sardinha, T. B. (2014) Comparing Internet and pre-Internet registers. In T. B. Sardinha and M. V. Pinto (eds.) *Multi Dimensional Analysis, 25 years on: A Tribute to Douglas Biber,* pp. 81—108. Amsterdam: John Benjamins Publishing Company

Sardinha, T. B. (2018) Dimensions of variation across Internet registers. *International Journal of Corpus Linguistics* 232: 125—157.

Sardinha, T. B. and Pinto, M. V. (2014) What's on TV? Looking at American television corpus linguistics style. Paper presented at the XVI Encontro de Alunos de Graduação em Inglês como Língua Estrangeira [XVI Meeting of Undergraduate Students in English as a Foreign Language], São Paulo, SP.

Sardinha, T. B., Kauffmann, C. and Acunzo, C. M. (2014) A multi-dimensional analysis of register variation in Brazilian Portuguese. *Corpora* 9(2): 239—271.

Sardinha, T. B., Pinto, M. V., Mayer, C., Zuppardi, M. C. and Kauffmann, C. H. (2019) Adding registers to a previous multi-dimensional. In T. B. Sardinha and M. V. Pinto (Eds.) *Multi-Dimensional Analysis: Research Methods and Current Issues*, pp. 165—186. London: Bloomsbury

Schegloff, E. A. (1992) "Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation". *American Journal of Sociology* 97(5): 1295—1345.

Schegloff, E. A. (2001) Getting serious: joke → serious 'no'. *Journal of Pragmatics* 33(12): 1947—1955.

Schler, J., Koppel, M., Argamon, S. and Pennebaker, J. (2006) Effects of Age and Gender on Blogging. Computational Approaches to Analyzing Weblogs. *Papers from the 2006 AAAI Spring Symposium, Technical Report SS-06-03*, Stanford, California, USA (March 27-29, 2006).

Schmidt, A. and Wiegand, M. (2017) A Survey on hate speech detection using natural language processing. In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*, pp. 1—10. Valencia, Spain.

Schmidt, J. (2011) (Micro)blogs: Practices of privacy management. In S. Trepte & L. Reinecke (Eds.) *Privacy online: Perspectives on privacy and self-disclosure in the social web,* pp. 159– 173. Heidelberg, Germany: Springer.

Schmidt, J. (2014) Twitter and the Rise of Personal Publics. In K. Weller, A. Bruns, J. Burgess, M. Mahrt and C. Pushmann (Eds.) *Twitter and Society*, pp. 221—236. New York: Peter Lang.

Schwartz, R., Tsur, O., Rappoport, A. and Koppel, M. (2013) Authorship attribution of micro-messages. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 1880—1891. Seattle, Washington, USA, (October).

Seargeant, P. and Tagg, C. (2014) *The Language of Social Media: Identity and Community on the Internet.* New York: Palgrave Macmillan

Selkie, E. M., Fales, J. L. and Moreno, M. A. (2016) Cyberbullying Prevalence Among US Middle and High School-Aged Adolescents: A Systematic Review and Quality Assessment. *Journal of Adolescent Health* 58(2): 125—133.

Sest, N. and March, E. (2017) Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences* 119: 69—72.

Shachaf, P. & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science* 36(3): 357—370.

Shih, G. (2019) Chinese censors go old school to clamp down on Twitter: A knock on the door. Retrieved on 8th August 2019 from https://www.washingtonpost.com/world/asia_pacific/chinese-censors-go-old-school-to-clamp-down-on-twitter-a-knock-on-the-door/2019/01/04/1bd462e4-f331-11e8-9240-e8028a62c722_story.html

Shoemark, P., Sur, D.. Shrimpton, L., Murray, I. and Goldwater, S. (2017) Aye or naw, whit dae ye hink? Scottish independence and linguistic identity on social media. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pp. 1239–1248, Valencia, Spain (April 3-7, 2017).

Shuang-Shuang, L. (2010) A Tentative Study of the Impoliteness Phenomenon in Computer-mediated Communication. *CROSS-CULTURAL COMMUNICATION* 6(1): 92—107.

Siegel, J., Dubrovsky, V., Kiesler, S. and McGuire, T. (1986) Group processes in computer-mediated communication. *Organizational Behavior and Human Decision Processes* 37(2): 157—187.

Singh, V., Varshney, A., Akhtar, S. S., Vijay, D. and Schrivastava, M. (2018). Aggression Detection on Social Media Text Using Deep Neural Networks. In *Proceedings of the Second Workshop on Abusive Language Online (ALW2)*, pp. 43–50 Brussels, Belgium (October 31, 2018).

Smith, B. (2003) The use of communication strategies in computer-mediated communication. System 31: 29—53.

Sourial, N., Wolfson, C., Zhu, B., Quail, J., Fletcher, J., Karunananthan, S., Bandeen-Roche, K., Béland, F. and Bergman, H. (2010) Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *Journal of Clinical Epidemiology* 63(6): 638—646.

Spencer, N. (2012) 1977: the Queen's punk jubilee. Retrieved on 8th August 2019 from https://www.theguardian.com/music/2012/apr/29/sex-pistols-queen-jubilee-boat

Spertus, E. (1997) Smokey: Automatic Recognition of Hostile Messages. In Proceedings of IAAI-97, pp. 1058—1065.

Stamatatos, E. (2009) A survey of modern authorship attribution methods. *Journal of the American Society for Information Science and Technology* 60(3): 538—556.

Stewart, I. and Eisenstein, J. (2018) Making "fetch" happen: The influence of social and linguistic context on nonstandard word growth and decline. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 4360–4370. Brussels, Belgium (October 31 - November 4, 2018).

Stieglitz, S. & Dang-Xuan, L. (2013) Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior. *Journal of Management Information Systems* 29(4): 217—248.

Sveningsson, M. (2014) "Not the real thing": Young people's experiences of getting civic and political news through social media. In *Proceedings of Culture, Technology and Communication*, pp.221—241. Norway: University of Oslo.

Synnott, J., Coulias, A. & Ioannou, M. (2017) Online trolling: The case of Madeleine McCann. *Computers in Human Behavior* 71: 70—78.

Tagliamonte, S. A. (2006) *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Tatman, R., Paullada, A., Stewart, L. G. and Spiro, E. S. (2017) Non-lexical features encode political affiliation on Twitter. In *Proceedings of the Second Workshop on Natural Language Processing and Computational Social Science*, pp. 63—67. (Vancouver, Canada)

Terrell, K. (2019) The History of Social Media: Social Networking Evolution! Retrieved on 8th August 2019 from https://historycooperative.org/the-history-of-social-media/

Thompson, P., Hunston, S., Murakami, A. & Vajn, D. (2017) Multi-dimensional analysis, text constellations, and interdisciplinary discourse. *International Journal of Corpus Linguistics* 22(2):153—186.
Thurlow, C., Lengel, L. and Tomic, A. (2004) *Computer Mediated Communication: Social Interaction and the Internet*. London: Sage.
Tiku, N. and Newton, C. (2015) Twitter CEO: "We suck at dealing with abuse". Retrieved on 8th August from https://www.theverge.com/2015/2/4/7982099/twitter-ceo-sent-memo-taking-personal-responsibility-for-the
Titak, A. and Roberson, A. (2013) Dimensions of web registers: an exploratory multi-dimensional comparison. *Corpora* 8(2): 235—260.
Tottie, G. (1983) *Much about not and nothing: a study of the variation between analytic and synthetic negation in contemporary American English*. Lund: CWK Gleerup
Trudgill, P. (1974) *The Social Differentiation of English in Norwich.* Cambridge: Cambridge University Press.
Tummers, J., Speelman, D. & Geeraerts, D. (2012) Multiple Correspondence Analysis as Heuristic Tool to Unveil Confounding Variables in Corpus Linguistics. In *Proceedings of the 11th International Conference on Statistical Analysis of Textual Data*, pp. 923—936. Liege, Belgium (June 13-15, 2016).
Twitter (2017) Authentication. Retrieved on 6 December 2016 from
Twitter (2018) The Twitter Rules. Retrieved on 24th January 2018 from https://help.twitter.com/en/rules-and-policies/twitter-rules
Twitter (2019a) homepage Retrieved on 6th June 2019 from https://twitter.com/
Twitter (2019b) homepage Retrieved on 14th September 2019 from https://twitter.com/
Uygur-Distexhe, D. (2012) Lol, mdr and ptdr: An inclusive and gradual approach to discourse markers. Lingvisticæ Investigationes 35(2).
Vaccari, C., Valeriani, A., Barberá, P., Jost, J. T., Nagler, J., & Tucker, J. A. (2016). Of Echo Chambers and Contrarian Clubs: Exposure to Political Disagreement Among German and Italian Users of Twitter. Social Media + Society 2(3): 1—24.
van Leeuwen, T. (2005) *Introducing Social Semiotics*. Oxon: Routledge.
van Rooy, B., & Terblanche, L. (2009). A multi-dimensional analysis of a learner corpus. In *A multi-dimensional analysis of a learner corpus*. Leiden, The Netherlands: Brill | Rodopi.
van Rooy, B., Terblanche, L., Haase, C. and Schmied, J. (2010) Register differentiation in East African English: A multidimensional study. *English World-Wide* 31(3): 311—349.
Vandebosch, H. and van Cleemput, K. (2008) Defining Cyberbullying: A Qualitative Research into the Perceptions of Youngsters. *CyberPsychology & Behavior* 11(4): 499—503.
Vandergriff, I. (2013) Emotive communication online: A contextual analysis of computer-mediated communication (CMC) cues. *Journal of Pragmatics* 51: 1—12.
Veszelszki, Á. (2017) Verbal and Visual Aggression in Trolling. In A. Benedek András, Á.Veszelszki (eds.) Virtual Reality − Real Visuality: Virtual, Visual, Veridical, pp. 141—151. Frankfurt am Main et al.: Peter Lang.

Viberg, Å. (2009). The verbs of perception: a typological study. *Linguistics* 21(1): 123—162.
Virtanen, T. (1992) Discourse Functions of Adverbial Placement in English: Clause-Initial Adverbials of Time and Place in Narratives and Procedural Place Descriptions. Åbo: Åbo Akademi University Press.
Wales, K. (2006) *Personal Pronouns in Present-Day English*. Cambridge: Cambridge University Press.
Wang, M. Y. and Silva, D. E. (2018) A slap or a jab: An experiment on viewing uncivil political discussions on facebook. *Computers in Human Behavior* 81: 73—83.
Warner, W. & Hirschberg, J. (2012). Detecting hate speech on the world wide web. In *Proceedings of the Second Workshop on Language in Social Media*, pp. 19—26. Montreal, Canada. (June 7, 2012).
Waseem, Z. & Hovy, D. (2016). Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter In *Proceedings of NAACL-HLT 2016*, pp. 88—93. San Diego, California. (June 12-17, 2016).
Waseem, Z., Thorne, J. and Bingel, J. (2018) Bridging the Gaps: Multi Task Learning for Domain Transfer of Hate Speech Detection. In J. Golbeck (ed.) *Online Harassment: Human-Computer Interaction Series,* pp. 29—55. Manhattan, N.Y.: Springer International Publishing.
Waseleski, C. (2006) Gender and the Use of Exclamation Points in Computer-Mediated Communication: An Analysis of Exclamations Posted to Two Electronic Discussion Lists. *Journal of Computer-Mediated Communication* 11(4): 1012—1024.
Weinreich, U., Labov, W., and Herzog, M. I. (1968) Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (eds.) *Directions for Historical Linguistics: A Symposium*, pp. 95—188. Austin: University of Texas Press.
Weinstein, E. C., Rundle, M. & James, C. (2015) A Hush Falls over the Crowd? Diminished Online Civic Expression Among Young Civic Actors. *International Journal of Communication* 9: 84—105.
Weller, K., Bruns, A., Burgess, J., Mahrt, M. and Puschmann, C. (2014) Twitter and Society: An Introduction. In K. Weller, A. Bruns, J. Burgess, M. Mahrt and C. Pushmann (Eds.)*Twitter and Society,* pp. xxix--xxxviii. New York: Peter Lang.
Wells, W., Burnett, J. and Moriarty, S. E. (2006) *Advertising: Principles and Practices*. 7th ed. Pearson.
Werry, C. C. (1996) Language and interactional features of Internet Relay Chat. In S. C. Herring (ed.) *Computer-Mediated Communication: Linguistic, social, and cross-cultural perspectives*, pp. 47—63. Amsterdam: John Benjamins Publishing Company
Westin, I. and Geisler, C. (2002) A multi-dimensional study of diachronic variation in British newspaper editorials. *ICAME Journal* 26: 133–152.
White M. (1994*).* Language in job interviews: Differences relating to success and socioeconomic variables. *PhD Dissertation*. Northern Arizona University.
Wikström, P. (2014) #srynotfunny. Communicative functions of hashtags on Twitter. *SKY Journal of Linguistics* 27: 127—152.

Wolf, A. (2000) Emotional Expression Online: Gender Differences in Emoticon Use. CyberPsychology and Behavior 3(5): 827—833.

Woolf, N. (2016) Lesley Jones bombarded with racist tweets after Ghostbusters opens. The Guardian (online). Retrieved on 6 September 2017 from https://www.theguardian.com/culture/2016/jul/18/leslie-jones-racist-tweets-ghostbusters

Wulczyn, E., Thain, N. & Dixon, L. (2017). Ex Machina: Personal Attacks Seen at Scale. In *Proceedings of the International World Wide Web Conference Committee* (IW3C2), pp. 1391—1399. Perth, Australia: Creative Commons.

Xiang, G., Fan, B., Wang, L., Hong, J. I., Rose, C. P. (2012) Detecting Offensive Tweets via Topical Feature Discovery over a Large Scale Twitter Corpus. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management* CIKM'12. Maui, HI, USA (October 29-November 2, 2012).

Xiao, R. (2009) Multidimensional analysis and the study of world Englishes. *World Englishes* 28(4): 421—450.

Xu, J., Jun, K., Zhu, X. and Bellmore, A. (2012) Learning from bullying traces in social media. In *Proceedings of the 2012 conference of the North American chapter of the association for computational linguistics: Human language technologies*, pp. 656–666. Montreal, Canada.

Yang, Y. and Eisenstein, J. (2017) Overcoming language variation in sentiment analysis with social attention. *Transactions of the Association for Computational Linguistics* 5: 295—307.

Yaqub, U., Chun, S. A., Atluri, V., and Vaidya, J. (2017) Analysis of political discourse on Twitter in the context of the 2016 US presidential elections. *Government Information Quarterly* 34: 613—26.

Yazdanfar, N. and Thomo, A. (2013) LINK RECOMMENDER: Collaborative-Filtering for Recommending URLs to Twitter Users. Procedia Computer Science 19: 412—419.

Yin, D., Xue, Z., Hong, L., Davison, B. D., Kontostathis, A. & Edwards, L. (2009) Detection of Harassment on Web 2.0" In *Proceedings of the Content Analysis in the WEB* 2, pp.1—7. Madrid, Spain. (April 21, 2009).

Yokota, K., Watanabe, K., Wachi, T., Otsuka, Y., Hirama, K. and Fujita, G. (2016) Crime linkage of sex offences in Japan by multiple correspondence analysis. *Journal of Investigative Psychology and Offender Profiling* 14(2): 109—119.

Zago, R. (2017) English in the traditional media: The case of colloquialisation between original films and remakes. In L. Landolfi, E. Federici and F. Cavaliere (eds.)*Transnational Subjects*: *Linguistic Encounters*. *Selected Papers from XXVII AIA Conference* volume II, pp. 91—104. Napoli: Liguori Editore.

Zappavigna, M. (2011) Ambient affiliation: A linguistic perspective on Twitter. *New Media and Society* 13(5): 788—806.

Zappavigna, M. (2015) Searchable talk: The linguistic functions of hashtags. *Social Semiotics* 25(3): 274—291.

Zappavigna, M. (2016) Social media photography: construing subjectivity in Instagram images. *Visual Communication* 15(3): 271—292.

Zappavigna, M. (2018). *Searchable talk: Hashtags and social media metadiscourse*. London: Bloomsbury.
Zhang, M. (2016). A multidimensional analysis of metadiscourse markers across written registers. *Discourse Studies* 18(2): 204—222.
Zhang, M., Sun, W., Peng, H., Gan, Q. and Yu. B (2017) A multidimensional analysis of metadiscourse markers across spoken registers. *Journal of Pragmatics* 117: 106—118.
Zhang, Z., Robinson, D. and Tepper, J. (2018) Detecting Hate Speech on Twitter Using a Convolution-GRU Based Deep Neural Network. In A. Gangemi, R. Navigli, M. E. Vidal, P. Hitzler, R. Troncy, L. Hollink, A. Tordai, and M. Alam (eds.) ESWC 2018: The semantic web. ESWC 2018. Lecture Notes in Computer Science, 10843, pp. 745-760 Springer Verlag.
Zhong, H., Li, H., Cinzia, A., Squicciarini, S., Rajtmajer, M., Griffin, C., Miller, D. J. and Caragea, C. (2016) Content-driven detection of cyberbullying on the in- stagram social network. In *IJCAI*, pp. 3952–3958. New York City, NY, USA. IJCAI/AAAI Press.
Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. *Nebraska Symposium on Motivation* 17: 237—307.
Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Oxford, England: Addison-Wesley Press.