Conflictual Interactions in CMC: the case of Twitter contexts

By

Xixiang Zhao

UCD Student Number: 17202801

The thesis is submitted to University College Dublin in fulfilment of the requirements for the degree of Doctor of Philosophy in Linguistics.

UCD School of Languages, Cultures and Linguistics

Head of School: Associate Professor Síofra Pierse

Principal Supervisor: Professor Bettina Migge

Members of the Research Studies Panel:

Associate Professor Máire Ní Chiosáin

Dr. Sandrine Peraldi

Dr. Sarah Kelly

October, 2022

Table of Contents

Abstract

In recent years, online conflict has received increasing attention from international organisations, government institutions and social media providers. It is typically treated as socially harmful, potentially undermining social stability and cyber security. International organisations (e.g., the Council of Europe, the European Union, the United Nations) have urged member nations to reach various international agreements on combatting online conflict. Government institutions (e.g., Cyprus, France, Germany) have passed laws and regulations to legislate against language use that incites social violence, hatred or discrimination. Social media companies (e.g., Twitter, Facebook, Instagram) have also established rules to restrict their users' language use and employ a significant number of staff to manually monitor interactions on their platforms. Unfortunately, this manual moderation of content has not substantially improved the verbal hygiene of the platform due to the constraints imposed by employees' workflow, workload and experience. In order to address this issue, data scientists have developed and optimised various kinds of computational conflict and hatred detection models which are mainly based on fixed linguistic features such as lexical, syntactic and semantic features. These detection models, however, do not appear to be highly effective at detecting online conflict that is characterised by language variation such as ever-changing Internet slang and non-orthographic spellings, and is highly interactionally dynamic in its use. There is thus an urgent need to explore online conflict by taking into account an interactional perspective.

However, most linguistic studies focus on the social issues and power imbalances that characterise conflictual interactions (e.g., Sahlane, 2019; Davies, 2019; Nahajec, 2019) and in terms of research on communication, only a few pragmatic studies are concerned with the discursive nature of conflict, especially in face-to-face contexts. They either aim to identify linguistic strategies in conflictual interactions (e.g., Shum & Lee, 2013; Ben-Menachem & Livnat, 2018) and examine the impact of contextual factors on the use of particular strategies (e.g., Kakavá, 2002; Fernández-Amaya, 2019) or explore the structure of conflictual interactions as a whole (i.e., how a conflictual interaction starts, develops and ends) (e.g., Bou-Franch & Blitvich, 2014). These studies do not address in detail how people's language use gives rise to and plays a role in the development of conflictual interactions though. To date, it is also not clear whether face-to-face conflictual interactions function in the same way as online ones because we still know too little about the latter. The present

study aims to contribute towards filling this gap by exploring the nature and developmental mechanisms of conflictual interactions in one social media context where conflict is prevalent, Twitter, by investigating the role of linguistic strategies and linguistic politeness in such interactions under the constraints and reshaping of Twitter algorithms.

The study examines the kinds of conflictual linguistic strategies Twitter users employ to perform conflictual interactions, their distribution and how these strategies drive the development of conflictual Twitter interactions. It also delves into the role of linguistic politeness in these strategies and in conflictual Twitter interactions. The analysis is conducted based on a corpus of conflictual Twitter interactions that includes 38 tweets and 11,188 replies with about 32,000 words. The in-depth textual analysis specifically designed corpus combines discourse analysis and the notion of facework. It focuses on how speakers design turns, how speakers' turns then trigger the next turns and finally how these turns are ordered chronologically to accomplish conflictual interactions with varying degrees of intensity. Given the sensitive nature of this analysis (i.e., researcher bias), the study employs two approaches: first, it pays close attention to the recipient's interpretation of the (previous) speaker's turn in order to gauge interlocutors' meanings, as suggested by discourse analysis. Second, two follow-up surveys are conducted to complement the textual analysis. The surveys focus on Twitter users' identification of conflictual linguistic strategies and their understanding of linguistic politeness, respectively. They assess the main findings of this study and provide additional information about language use that is not easily available from the textual analysis alone.

Six main findings emerge from the analysis: first, like face-to-face conversations, participants' roles in conflictual Twitter interactions are also constantly changing, alternating between being a recipient and a speaker, to issue the next turn. As current speakers, their turns are based on their understanding of the previous speaker's turn and their own agendas and on other users' (public) interpretations of their turns. When a new turn is published, they finally become the previous speaker of that new turn. By successively carrying out this role shift, Twitter users cooperatively achieve conflictual interactions. Conflictual interactions continually evolve by users taking turns posting new turns in the same way as the current speaker. Second, conflictual interactions typically arise on Twitter when Twitter users disagree with other users' viewpoints or detest other users. As the interaction progresses, its intensity then gradually escalates and suddenly drops to zero at some point

when one of the participants withdraws from the interaction. Third, in such interactions, the current speaker objects to the previous speaker's point of view using disagreements and/or personal attacks. Around one third of the total posts in the corpus contain personal attacks which are generally considered taboo in face-to-face contexts. Fourth, the literal function of Twitter users' language use may differ from its interactional function: not all posts taken as a public attack by the recipient constitute personal attacks from a structural point of view; not all posts containing personal attacks are responded to with attacks. Such mismatches occur for various reasons: the recipient may have considered the (previous) speaker's attack less offensive or may have just decided to ignore the (previous) speaker's attack to pursue their own interactional goals. Fifth, when the recipient believes that they are being attacked and determines to attack the (previous) speaker, in return employing personal attacks, they tend to normalise personal attacks as an appropriate and effective strategy and to use attacks of a more aggressive nature to ensure that they are heard and that their anger is registered. This results in conflictual interactions being prolonged and escalating. Finally, almost all of the total posts (including both disagreements and attacks) in the corpus include linguistic politeness which functions to modify (i.e., mitigate or intensify) face-threatening acts. Although linguistic politeness is mostly used to mitigate the possible threat to the recipient's face, more than one third of the total posts in the corpus including those involving linguistic politeness features still pose serious damage to the recipient's face and are taken as an overt attack by the recipient.

Three major conclusions are drawn based on these findings: first, linguists should pay equal attention to both conflictual interactions and harmonious ones because both are an integral part of human communication and both are a result of participants' cooperation. Second, changes in people's uses and perceptions of linguistic strategies on Twitter suggest that technology appears to participate in, constrain and shape human interactions and language use in virtual settings. For example, personal attacks, which are considered disadvantageous and worthless in face-to-face contexts, seem to serve as an acceptable linguistic strategy on Twitter to register speakers' anger, to announce their dominance of opinion and to drive the development of interactions. Linguistic politeness, which is commonly regarded as a strategy to preserve verbal hygiene in face-to-face conversations, is however also used by the speaker to intensify the damage to the recipient's face in conflictual Twitter interactions. This suggests that in computer-mediated communication

studies, the role of technology needs to be carefully considered. Third, the study also highlights the important role of interactivity and the affordances of interactions in understanding and analysing conflictual interactions. It also argues that this fact may be responsible for the failure of feature-based computational conflict detection models as they typically focus on isolated turns or units with turns only rather than entire interactions. The analysis shows that attacks are not isolated instances but arise out of an interaction and perform clear discursive functions within the interaction such as ensuring that the speaker is heard. It is also worth noting that the speaker's attacks do not necessarily trigger the recipient's attacks and that giving up attacking does not guarantee that the speaker would not be attacked by the recipient either. The interpretation of the (previous) turn and construction of the next turn are fully the choice of the recipient. The analysis suggests that the recipient's actions—when they read and respond to the previous turn, determine how to interpret its strategic design and communicative intent, how to decide their own strategic goals and how to construct their response—have a greater impact on conflict than the (previous) speaker's language use. Their actions bridge the previous and the next turns; their interpretation and strategic goals make visible the function that the (previous) speaker's turn really plays in an interaction.

Keywords: conflictual interactions, computer-mediated communication, technology, Twitter, discourse analysis, pragmatic strategy, linguistic politeness

Dedicate to my mum
and
those good old days

Acknowledgement

I am eternally grateful to all people who have helped and inspired me during my doctoral study. This thesis would have not been possible without the help, support, guidance and encouragement of many people.

First and foremost, I would like to express my heartfelt gratitude and sincere appreciation to my supervisor Professor Bettina Migge, for all of the expertise, guidance and support she provided me throughout each stage of this dissertation project. I am greatly indebted to her for the trust she had in me when accepting me as her Ph.D. student and giving me the opportunity to carry out this research, which was not generally well received when it first began. It would have never been possible for me to complete this thesis without her continuous encouragement, immense knowledge, insightful advice and invaluable comments. I also want to thank Bettina for her patience to proofread all drafts of this thesis regarding issues such as English language uses and academic writing styles, which was extremely important to me as a non-native English speaker. Bettina was the most understanding, patient and intelligent person throughout my entire graduate studies at the University College Dublin, which enabled me to overcome setbacks and to become a more confident, self-disciplined and creative person; and for that, I'll always be grateful for her support.

I also want to extend a very special thanks to other members in my research panel, Associate Professor Máire Ní Chiosáin, Dr. Sandrine Peraldi and Dr. Sarah Kelly, for their support and insights during the writing of this dissertation. Very special thanks are due to Sarah Kelly for her consistent interest in this study, her constructive comments and inspiring discussions regarding many aspects of this dissertation, her expertise and instructive support when I was designing the questionnaire surveys and her meticulous review of individual chapters and full drafts of the thesis during the final phase. She was very generous with her time and knowledge in the process of completing the thesis. Thanks too to all the staff and Ph.D. students based in the School of Languages, Cultures and Linguistics at the University College Dublin for their much-needed encouragement.

My thanks also go out to my family and friends for providing me with unwavering support, immense love and continuous encouragement. This thesis is dedicated to my mum, a wonderful lady and the most important person in my life. I would also like to extend my

heartfelt thanks to my best friends Qiu, Yao and Lian, who put their trust in me all these years and stood up for me in difficult times. Very special thanks are due to Dr. Deng, Jiahan and Rui who always listened and provided valuable ideas when I was stuck in my thoughts during the writing phase of this dissertation. I really hope that I have done you all proud.

I would especially like to thank my landlords Claire and Conor McWade, who made me feel at home in Dublin and provided generous help when I was sick, robbed and scammed. A special thanks is also due to my neighbour's orange cat for guarding me on my way home at night. He was always able to cure my tiredness and anxiety of the day by running towards me, rubbing against my legs and purring.

I also want to express my appreciation to the China Scholar Council for granting me a scholarship, which enabled me to pursue my doctoral degree. Finally, thanks to all those others not mentioned here but who supported and encouraged me.

Statement of Original Authorship

I hereby certify that the submitted work is my own work, has completed while registered as a candidate for the degree stated on the Title Page, and I have not obtained a degree elsewhere on the basis of the research presented in this submitted work

List of Tables

List of Figures

List of Special Terms and Abbreviations

AP: adjacency pair

API: Application Programming Interface

BOW: Bag of Words

CA: conversation analysis

CMC: computer-mediated communication

CONVERSATION: an interaction including an original tweet and all posts triggered by that tweet;

CP: cooperative principle

DA: discourse analysis

ECRI: European Commission against Racism and Intolerance

EU: European Union

FPP: 1st pair part

FTA: face-threatening act

GIF: graphic interchange format

Level: indicating the distance from the original tweet. For example, 1st level reply means that the reply is the direct comment on the tweet; 2nd level reply means that there is another reply between that reply and the orginial tweet.

MP: model person

NLP: natural language processing

One-on-one interaction: the longitudinal interaction that includes an original tweet and a string of corresponding upper-level replies from different levels.

One-to-many interaction: the latitudinal interaction that consists of a reply or an original tweet with its corresponding upper-level replies at the same level.

POS: Part-of-Speech

RADAR: Regulating Anti-Discrimination and Anti-Racism

SPP: 2nd pair part

Stream: an one-on-one interaction that derives from an original tweet and ends with no more replies.

TCU: turn-constructional unit

TRP: transition-relevance place

Turn: all replies published by a Twitter user in response to the previous speaker;

UN: United Union

Chapter 1 Introduction

1.1 Introduction to the subject matter

Social conflict pertains to "any situation or behaviour involving parties (individuals or groups) who are or consider themselves to be, instrumentally, intellectually and/or emotionally opposed or simply feel antagonistic to each other" (Jeffries & O'Driscoll, 2019, p. 2). It is an ordinary phenomenon in everyday life, occurring during any controversial activity or event such as an argument, fight, quarrel, debate, dispute, bickering, altercation, contention, dissension, quibbling, squabble and wrangle and within any situation such as an arbitration, family dinner, hospital, school lunch, theatre rehearsal, workshop and netsurfing (Grimshaw, 1990). Karl Marx developed credible and powerful analyses of conflict between social classes in the 19th century, which initiated conflict studies in social science (Bartos & Wehr, 2002). Since then, a growing number of social scientists (e.g., sociologists, political scientists, anthropologists) have shown interest in conflict studies, seeking to explain and solve increasingly serious social problems. They devoted considerable effort, for instance, to investigating the social and psychological causes of cyberbullying, detecting hateful language against vulnerable groups in society and modelling the development of social conflict.

Communication problems are generally at the root of all types of conflict, although the underlying causes of conflict are rarely linguistic. As early as the end of the 20th century, a handful of linguists were aware of this fact (e.g., Grimshaw, 1990; Briggs, 1996; Pearce & Littlejohn, 1997) and coined the term 'conflictual interaction' to describe the discursive nature of conflict. *Conflictual interaction* refers to an interaction in which participants take alternative positions on an issue of common concern (e.g., Putnam & Poole, 1987; Grimshaw, 1990; Vuchinich, 1990; Hutchby, 1992; Putnam, 2001; Honda, 2002; Bardovi-Harlig & Salabury, 2004; Norrick & Spitz, 2008; Zhu, 2008; Nguyen, 2011; Chi, 2016; Koester, 2017; Blitvich & Bou-Franch, 2019; O'Driscoll, 2019; Schmitt & Reiter, 2019). A growing number of applied linguists have paid attention to this type of interaction in recent years and view it as a static and pre-existing social state manifested through language (e.g., Sahlane, 2019; Davies, 2019; Nahajec, 2019; Alaghbary, 2019; Blitvich, 2019; Khan, 2019; De Bres & Franziskus, 2019; Archer et al., 2019; Coinnigh et al., 2019). Harrison et al.'s (2019) study of fish stocking is a case in point. Fish stocking is an activity aimed at "enhancing fishing opportunities, compensating for degraded environments, replacing missing or

dysfunctional spawning sites and supporting threatened or declining populations" (Harrison et al., 2019, p. 508). However, it generates heated debates among stakeholders regarding its risks and benefits from a socioeconomic and ecological perspective. In order to gain insights into the broader social underpinnings of such debates, Harrison et al. (2019) collected multi-faceted data about the termination of Atlantic salmon stocking in Wales in 2014 (e.g., interviews, news articles, Internet postings, official documents) and provided guidance on the sustainable management of wild fish populations that takes account of social and ecological impacts. Applied linguists, like Harrison et al. (2019), are particularly concerned with revealing the ideology or imbalance of social power that underlies social conflict and proposing possible solutions. They merely focus on "the role of languages and dialects themselves in decisions relating to conflict" from a socio-political standpoint and on the "ideational features of texts" from a textual perspective (Jeffries & O'Driscoll, 2019, p. 8) because they generally hold that investigating the discursive nature of conflict is less practical than mitigating or managing social issues. This results in this type of linguistic research contributing little to the understanding of conflict as it occurs in action.

The discursive nature of conflict is marginalised in modern linguistic research (especially in pragmatics, socially oriented pragmatics and interactional sociolinguistics) on interactions, except for a few studies on disagreement strategies, such as Schiffrin (1985), Goodwin and Goodwin (1987), Johnstone, (1989) and Brown (1990) (which will be discussed in detail in Chapter 2). There are two possible reasons. First, the Gricean Cooperative Principles (CPs) (1975), a cornerstone of most interactional theories or models, are commonly misinterpreted as meaning that "in their interactions people always agree with each other or…[are] nice to each other" (Sifianou, 2019, p. 187). This misunderstanding is a default interpretation based on how individuals interpret others and/or their own contributions in an interaction. Grice, however, did not provide any definition of cooperation or assert that harmonious interactions necessarily lead to cooperation or deny that cooperation is possible in conflictual interactions. He even argued that "within the dimension of voluntary exchanges…collaboration is achieving exchange of information, or the institution of decisions may coexist with a high degree of reserve, hostility and chicanery" (Grice, 1989, p. 369). Second, it may also be due to the Anglocentric nature of much of the research on pragmatics and to the reliance on English-based data (Wierzbicka,

2011; Levisen, 2016, 2019).[1] Researchers in interactional studies generally hold "fairly Anglo and gentlemanly norms and beliefs" (Levisen, 2019, p. 2), viewing linguistic politeness as typical of English discourse (Klein, 2009). Their preference for interpersonal harmony is subsequently reinforced by a series of Anglosphere based theories of linguistic politeness, most notably Brown and Levinson's (1987/1978) face-saving theory. Despite a growing concern over linguistic impoliteness since the 2000s (e.g., Bousfield, 2008; Locher & Watts, 2005, 2008; Locher & Bousfield, 2018; Haugh & Culpeper, 2018; Culpeper, 2021; Haugh, 2022), that is a focus on how the speaker maximises the damage to the recipient's face, core issues such as terminological distinctions and methodological matters are still under heated discussion. As a result, compared to studies on harmonious interactions, only a few interactional studies have systematically analysed conflictual interactions involving inharmonious language use in detail. It was not until the prevalence of virtual contexts that the issue of conflictual interactions was formally brought to the table of interactional research. This significant shift is closely related to three facts brought about by the development of Internet technologies.

First, with the widespread use of online interactions, especially on social media such as Twitter, what would be considered uninhibited linguistic behaviour in face-to-face interactions has come to pose a growing threat to cybersecurity and social stability and has become an urgent issue to be addressed (e.g., Murthy, 2013; European Commission against Racism and Intolerance [ECRI], 2015; Regulating Anti-Discrimination and Anti-Racism [RADAR], 2015; European Union [EU], 2016; Siapera et al., 2018; Graham, 2019; Haugh & Sinkeviciute, 2019; Sifianou, 2019; Salminen et al., 2020; United Union [UN], 2020). Although they are not unique to virtual contexts (e.g., Council of Europe, 1997; Georgakopoulou, 2001; Kakavá, 2002; Ben-Menachem & Livnat, 2018; UN, 2020), the lack of established norms for these newly emerging interactive contexts has allowed them to flourish in these settings and has led to them becoming stereotypically associated with these environments. The Internet thus provides an appropriate setting for studying inharmonious language use because not only is uninhibited linguistic behaviour more common in virtual environments, but it is also easier for researchers to obtain a record of conflictual interactions. This contrasts with the case in face-to-face contexts where the occurrence of conflictual

[1] Anglocentrism refers to a "tacit believe in the nature and neutral status of English words and meanings and a lack of attention to the way in which such words shape and form research, results and claims we make about humanity at large" (Levisen, 2019, p. 1).

discourses is difficult to detect and participants in such discourses may not be willing to be recorded. Despite this stereotype and convenience, a lot of linguists still tend to base their analyses of uninhibited linguistic behaviour in face-to-face interactions (e.g., Muntigl & Turnbull, 1998; Georgakopoulou, 2001; Kakavá, 2002; Ben-Menachem & Livnat, 2018). Little is known about the linguistic nature and interactional functions of inharmonious language use on the Internet.

Second, to improve the verbal hygiene in virtual contexts, data scientists have developed various classifiers for the computational detection of inharmonious language use on the Internet (e.g., Davidson et al., 2017; Mondal et al., 2017; Saleem et al., 2017; Sahlgren et al., 2018; Salminen et al., 2018ab). Currently, the most popular classifiers are based on textual features such as keywords, Bag of Words (BOW), N-grams and Part-of-Speech (POS), considering online conflict as a set of fixed linguistic features with identical functions.[2][3][4] However, these feature-based classifiers have worked poorly in practice (Watanabe et al., 2018) because the meaning and function of words constantly change when used across different contexts. For example, the hedge *please* is not necessarily always polite because in the phrase 'Can you fuck off please?', it functions to intensify the offence. While the address *little bitch* is generally taken as an insult or a negative reference to someone, it sometimes serves as a term of endearment between lovers.[5] Even though researchers constantly update the language features for the classifiers, users can easily evade computational detection by creating new language forms or patterns. They seem to play a game of peek-a-boo with the classifiers because they are always able to create novel linguistic variants after certain expressions have been banned by the algorithm. For example, in cases of the term *fuck*, the most common way is to replace the middle letters *u* and *c* with multiple asterisks (e.g., *f*****k*). People sometimes even use a graphic interchange format (GIF) that displays no textual information but only a mouth that utters the word *fuck* without sound. These

[2] Bag of Words refers to "a counting algorithm that encodes a sentence or a document in the dataset using a vector representation with |V| dimensions where |V| is the size of the selected vocabulary, V. The representation is unordered and encodes the weight of each member of the V, in the given instance, using various scoring techniques" (Salminen et al., 2020, p. 16).

[3] N-grams refer to "a substring of length n characters derived from a text string (usually but not necessarily, a word) containing not less than n characters. The characters in the n-gram retain the same order as in the source text from which the n-gram has been derived" (Robertson & Willett, 1998, p. 49).

[4] Part-of-Speech tagging refers to "the process of assigning appropriate lexical category to [an] individual word in a sentence of a natural language" (Patra et al., 2012, p. 923).

[5] On Quora, a social question-and-answer website, the question 'Do you like being called a bitch during sex?' received answers from four female users. They provided affirmative answers, arguing that being called a bitch during sex greatly simulates their sexual desire and pleasure. For more information, see: https://www.quora.com/Do-you-like-being-called-a-bitch-during-sex/

issues suggest that linguistic features are less effective in identifying online conflict that is highly dynamic and contextualised and prompt researchers to reconsider the value of studying the dynamic nature of interactions and how it functions in virtual contexts. There are a number of (im)politeness studies that investigate the nature of this dynamic (e.g., Leech, 2004/1983; Culpeper, 2011ab; Haugh, 2022), but these linguistic findings from face-to-face and/or online contexts have not been sufficiently applied to computational techniques for improving the accuracy of online conflict management.

Third, linguists consider humans to be the only subjects capable of affecting the nature, evolution and development of online interactions, which are fundamentally similar to face-to-face interactions (e.g., Androutsopoulos, 2006; Chua et al., 2010; Shum & Lee, 2013; Xie et al., 2013; RADAR, 2015; Tagg, 2015; Fernández-Amaya, 2019; Ilbury, 2020). They are concerned with the ability of individuals to use language in online conversations. Specifically, linguists examine how people employ language to interact with others on the Internet, how contexts (e.g., real/virtual identities, technical issues, cultural factors, ethical standards, interpersonal relationships) affect the ways in which people use language and/or interact and how language use and/or interactional patterns in turn shape people's personal images or status in virtual settings. This suggests that online interactional studies still follow a humanist perspective, viewing interpersonal interactions and mutual understanding as the sole focus (also see Pennycook, 2017). However, technology is actively engaging in human communication on the Internet. It redefines the rules of human communication. For example, people in online interactions are unable to use or capture many of the informational cues that are quite common in face-to-face interactions such as gestures, postures, facial expressions, intonation, social context information, social feedback and personal details (e.g., clothes, social status, occupation, age, gender, name; e.g., Baron, 1984; Siegel et al., 1986; Myers, 1987; Reid, 1991; Aycock, 1995; Baym, 1995; Lee, 1996; Parks & Floyd, 1996; Postmes et al., 1999). Their communication is also governed by the interactional rules set by the algorithms of different online platforms. A case in point is Facebook, which requires that participants have to be each other's followers. On Twitter, in contrast, interactions between strangers are possible because users can view posts from another user they follow and retweet or comment on that user's posts without asking for permission. Technology also changes people's ideologies of language use and communication. Language ideology refers to "a totalising vision" and/or "a regime of value"

that can be used to account for people's ordinary language practices in everyday life (Gal & Irvine, 2019, p. 13). For example, Herring (1996a) found that it is impossible to distinguish the language used on the Internet based on the classical dichotomy that divides natural language into spoken and written language. Although online language has similarities to written language because it displays as texts on screen in most cases, it also resembles spoken language; it contains multimodal features (e.g., emoticons, emoji, acronyms) and is loosely constrained by the traditional written norms. In terms of communication, a lot of online platforms provide in-built tools that contribute to both synchronous and asynchronous interactions, such as iMessage, WeChat, WhatsApp, Facebook and Twitter, thus challenging the stereotype that human communication is either synchronous or asynchronous. In online contexts, people are able to communicate synchronously by "giv[ing] or receiv[ing] an immediate response" (Donelan, 2010, p. 86) and/or to develop asynchronous interactions by delaying their responses. They can also keep track of their conversations at any time by using the function of diachronic threads to read the trail of previous messages (e.g., Wallace, 1999; Donelan, 2010). All of these changes indicate that technology, as an extension of humans and their sense, is actively involved in shaping human life (McLuhan, 1964). It "weave[s] into the fabric of [human] language" and "reshape[s] the nature of face-to-face interaction[s]" (European Cooperation in Science and Technology, 2021, p. 6). It also "hold[s] complex and lengthy conversations with [humans]" as an "agent of communicative acts", developing new interactional rules and creating original interactional patterns (European Cooperation in Science and Technology, 2021, p. 7). As Pennycook's (2017, p. 14) posthumanist perspective maintains, technology "questions human minds as central to… action and interaction and questions the distinctions between humans and… object". Interactions on the Internet are no longer restricted to human-human interactions but also include human-machine interactions which have received little attention from linguists.

To contribute towards filling the above-mentioned gaps (i.e., little is known about the discursive nature of conflict, the linguistic nature of online conversations and the role of technology in shaping human communication and language use), the present study focuses on the relationship between language use and the dynamics of online conflictual interactions. It pays particular attention to Twitter contexts as according to a survey of multiple online platforms conducted by the EU in 2019, uninhibited linguistic behaviour is

particularly insensitive and frequent on Twitter. This study adopts a multidisciplinary and posthumanist approach: it considers technology as a social factor that impacts humans' language use and interactions, conducts a sociolinguistics-oriented pragmatic analysis of Twitter interactions and takes into account some related theories and methods from the fields of computer-mediated communication, corpus linguistics and natural language processing. It aims to explore the linguistic mechanisms that give rise to the occurrence and development of conflictual interactions on Twitter and to delve into the specific ways in which people engage in conflictual Twitter interactions under the constraints and reshaping of the Twitter algorithm.

1.2 Ubiquitous conflictual interactions in CMC

This section sets out the definition and scope of computer-mediated communication (CMC) to set the scene for investigating conflictual interactions in virtual contexts. It then surveys how CMC alters human interactions and explores possible reasons for the prevalence of conflictual interactions in online environments, especially on Twitter.

1.2.1 What is CMC?

As an unplanned by-product of computer technologies, CMC originally referred to the transmission of data (e.g., security, information redundancy) between large computers (Walther, 1996). Upon realising that CMC transmits data as well as simple messages, researchers hypothesised that it might be possible to overcome geographic inconvenience and to coordinate emergency tasks via CMC (Walther, 1996). This led to CMC becoming a planned application for group communication and eventually stepping into human life (Walther, 1996). Human language use in offline contexts (e.g., face-to-face interactions, academic writing, letters) is thus extended to virtual contexts. This change in human interactions has greatly stimulated research interests in online linguistic behaviour. A new interdisciplinary field of CMC was eventually established by Susan C. Herring in the late 1990s.

Herring (1996a, p. 1) defined *computer-mediated communication* as "communication that takes place between human beings via the instrumentality of computer". Her definition has become a classic and has profoundly influenced further CMC studies. According to Herring (1996a), CMC belongs to human communication and is carried out by computational

media. Many researchers have since developed Herring's definition in different ways. December's (1997) definition, for example, increased the focus on the circumstances of CMC. He defined CMC as "a process of human communication via computers, involving people, situated in particular contexts, engaging in processes to shape media for a variety of purpose[s]".[6] Thurlow et al. (2004, p. 15) regarded CMC media as a developing technology (i.e., computer technology) rather than an instrument (i.e., computer) and referred to CMC as "any human communication through or with the help of computer technology". McArthur et al.'s (2018/2013) definition in *The Oxford Companion to the English Language (2nd Edition. online version)* highlighted the interactive nature of CMC. It takes CMC as a "human-human interaction mediated via electronic devices", breaking the restriction on the nature and modes of CMC.[7] This study considers CMC as a broad description of all types of human interactions that occur in various electronic devices (e.g., laptops, mobile phones, tablets, PCs, Apple watches), including verbal and other visual or aural communication (e.g., sending emoji, pictures and voice mails; sharing music and videos).

1.2.2 Impact of CMC on human interactions

Early CMC research generally held negative views about CMC (e.g., Daft & Lengel, 1983; Baron, 1984; Kiesler et al., 1984, 1985; Rice, 1984; Siegel et al., 1986; Myers, 1987; Dubrovsky et al., 1990; Reid, 1991; Sproull & Kiesler, 1991; Garton & Wellman, 1995; Aycock, 1995; Baym, 1995; Lee, 1996; Parks & Floyd, 1996; Postmes et al., 1999). Researchers believed that the absence of informational cues in CMC (e.g., gestures, postures, facial expressions, intonation, social context information, social feedback and personal information such as clothing, social status, occupation, age, gender and name) reduces the expressiveness and complexity of natural language. It becomes quite challenging for users to figure out what is happening and what are the real meanings and implications of textual information. Such negative views were probably established because most of the early studies were carried out prior to the extensive commercialisation of the Internet in the mid-1990s. They seldom addressed web applications (e.g., Facebook, Instagram, Twitter) that allow users to share information, to interact and to cooperate with others virtually. Their

[6] Accessed, 22nd February 2022: https://www.december.com/cmc/mag/1997/jan/december.html
[7] Accessed, 22nd February 2022: https://www.oxfordreference.com/view/10.1093/acref/9780199661282.001.0001/acref-9780199661282-e-1363?rskey=MuCUxH&result=310

results are, thus, more appropriate for text-based media, rather than multimodal and synchronous communication media.

Innovations in technology, however, have not erased the fact that the lack of personal information and social cues leads to deindividuation in CMC (Reid, 1991). CMC allows users to disclose or display their real identities to suit their communication needs. Not only are they able to display their online presence by disclosing some or all of their offline identities but they can also hide or manipulate their offline identities in online contexts to minimise the impacts of their online behaviour on their reputation in the real world (Reid, 1991). Thus, users are left with a consistent impression of fragmentation between the virtual and the real created by CMC.

Deindividuation diminishes the impact of higher-status perspectives and reduces bias and prejudice against people of low-status (e.g., Kiesler et al., 1984; Siegel et al., 1986; Myers, 1987; Reid, 1991; Sproull & Kiesler, 1991; Baym, 1995; Garton & Wellman, 1995; Aycock, 1995; Lee, 1996; Parks & Floyd, 1996). It provides individuals with a relatively equal right to communicate with more confidence and challenges the traditional hierarchical systems and status barriers. Postmes et al.'s (1999) work demonstrates that deindividuation within small groups helps members to develop situational norms and to foster a sense of belonging through in-group interactions. On the other hand, deindividuation causes more frequent uninhibited behaviour (e.g., verbal aggression, blunt disclosure, non-conforming behaviour) in CMC than in face-to-face interactions (e.g., Kiesler et al., 1984; Siegel et al., 1986; Reid, 1991; Sproull & Kiesler, 1991; Garton & Wellman, 1995; Aycock, 1995; Parks & Floyd, 1996; Postmes et al., 1999). By challenging social norms, which refer to a "specific guide to action which defines acceptable and appropriate behaviour in particular situations" (Haralambos & Holborn, 1991/1980, p. 5), uninhibited behaviour greatly reduces people's feelings of embarrassment, guilt and sympathy for others as well as their fear of retribution and rejection. People thus behave more rudely in CMC and generate much more conflictual interactions on the Internet than in face-to-face settings. For example, in Siegel et al.'s (1986) research, there were 54 student computer-mail users reporting 34 instances of swearing, insults and name-calling while none occurred in face-to-face interactions. Dubrovsky et al. (1990) also found that students became less efficient and more uninhibited when using email than in face-to-face communication.

In recent years, the problem of deindividuation has become increasingly severe with the widespread use of social media (Murthy, 2013; Haugh & Sinkeviciute, 2019; Graham, 2019; Sifianou, 2019). CMC serves as an extended space for real-life social issues. One of the most typical phenomena is the expression or evocation of social hatred against certain ethnic or other vulnerable groups on the Internet. The language used in this case is what researchers call *hate speech*, which becomes a major threat to social media (e.g., Sood et al., 2012; Dinakar et al., 2012; Chen et al., 2012; Dadvar et al., 2013; Kwok & Wang, 2013; Gitari et al., 2015; Burnap & Williams, 2015; Kansara & Shekokar, 2015; Djuric et al., 2015; RADAR, 2015; ECRI, 2015; Nobata et al., 2016; Mondal et al., 2017; Wulczyn et al., 2017; Davidson et al., 2017; Park & Fung, 2017; Adamczak-Krysztofowicz & Szczepaniak-Kozak, 2017; Assimakopoulos et al., 2017; Salminen et al., 2018ab, 2000; ElSherief et al., 2018; Watanabe et al., 2018; Zhang et al., 2018; Siapera et al., 2018; UN, 2020). By deliberately violating socially and culturally accepted norms in face-to-face communication, CMC users who share a similar hatred of a vulnerable group naturally congregate. They aim to establish "a hierarchy of difference among minorities" and to create "a relation between the majority and minority groups based on features that [the users] inherited and thus cannot change" (Adamczak-Krysztofowicz & Szczepaniak-Kozak, 2017, p. 289). Hate speech undermines social cohesion, erodes democratic values and threatens the rule of law and human rights. It may further destabilise democratic societies, sabotage social peace, lay a foundation for violence and increase the possibility of alienation and radicalisation. In 2015, the ECRI stated that if left unaddressed, the spread of hate speech online may result in much greater violence and conflict, thereby seriously undermining social cohesion.[8]

Twitter, a CMC platform with the most pervasive conflictual interactions, was created in March 2006, with over 330 million monthly active users and 500 million daily tweets by April 2019 (Statista Research Department, 2019). 21% of users employ Twitter for conversations and 12.5% of tweets are part of conversations (Java et al., 2007). Although up to date, as far as I know, no statistical data is available to demonstrate that conflictual interactions on Twitter are much more prevalent than on other CMC modes, a number of reports, literary works and lay and professional discourses closely associated Twitter with

[8] In 1993, the Council of Europe established the European Commission against Racism and Intolerance (ECRI), which aims to combat racism, discrimination (on grounds of race, ethnic/national origin, color, citizenship, religion, language, sexual orientation, gender identity and sex characteristics), xenophobia, antisemitism and intolerance. The main task of ECRI is to prepare reports on hate speech in member states and to issue recommendations of legislation on hate speech to member states.

conflict (e.g., Peddinti et al., 2017; Sterner & Felmlee, 2017; Jonathan, 2017; EU, 2019; James, 2019; Prostko, 2020). The overall design of Twitter also promotes deindividuation and uninhibited linguistic behaviour. For example, according to Ilbury (2020), in order to eliminate the influence of real life and to escape the restrictions of their real identities, Twitter users may hide their real identities according to their communication needs, and some of them may even create new ones. Given the protection of personal privacy, Twitter also protects these actions and allows anonymity. It does not force users to provide real identities, nor does it allow researchers to associate users' accounts with off-Twitter identifiers without permission. As a result, uninhibited linguistic behaviour (e.g., verbal attacks, outspoken disclosure, unruly behaviour) becomes particularly prevalent on Twitter.

Collectively, the review suggests that CMC is a double-edged sword. A good deal of Informational cues which are quite common in face-to-face conversations, such as gestures, facial expressions, intonation, social context information, social feedback become scarce in online contexts, and are replaced with reduced types of feedback (e.g., nods, murmurs, frowns, eye contact), resulting in the deindividuation of individual communication. On the one hand, deindividuation challenges traditional hierarchical systems and status barriers and encourages equal communication. CMC users are able to express their own opinions without fear of repercussions on their offline identities, which promotes online democracy. On the other hand, possibly more importantly, deindividuation causes uninhibited linguistic behaviour (e.g., verbal aggression, blunt disclosure, non-conforming behaviour) and results in the prevalence of conflictual interactions in CMC, especially on Twitter. The next section continues to investigate the responses of different social areas such as public policies, data scientists, the Twitter company and linguistics, to the prevalence of conflictual interactions in CMC.

1.3 Concerns about and responses to conflictual interactions

This section examines the measures taken by the relevant social institutions and professionals and the Twitter company in response to the prevalence of conflictual interactions in CMC. It briefly explores the current dilemmas faced by automatic conflict detection techniques, preliminarily surveys linguistic research on conflictual interactions and identifies both the potential academic and social areas in which this study can make an impact.

1.3.1 Responses from the public

Online conflictual interactions involve uninhibited linguistic behaviour used to express hatred, such as hate speech and also behaviour used to communicate opposing viewpoints, such as arguments, disputes and disagreements. However, only hate speech has been taken seriously by the public. As outlined in the *Recommendation R (97) 20 of the Committee of Ministers to Member States on "Hate Speech"* (hereinafter referred to as *Recommendation R (97) 20*), member governments were advised to "take appropriate steps to combat hate speech", particularly, the "hate speech disseminated through media" (Council of Europe, 1997, pp. 107–108). In the *ECRI General Policy Recommendation N°15 on Combating Hate Speech*, the ECRI (2015) characterised hate speech as an extreme form of intolerance associated with violence. To redress the pernicious impact of the spread of hate speech on social media, in 2016, the EU and IT companies (i.e., Facebook, Microsoft, Twitter and YouTube) agreed on the *Code of Conduct on Countering Illegal Hate Speech Online* (hereinafter referred to as *Code*). According to the *Code*, the IT Companies have the responsibility to share "the European Commission's and EU Member States' commitment to tackle illegal hate speech online" (EU, 2016, p. 1). For example, upon receipt of valid removal notifications, they should review most notifications within 24 hours. If necessary, they are also supposed to remove or disable access to the content in accordance with its rules or community guidelines. In France, in 2019, Laetitia Avia, a National Assembly member, proposed a bill for combatting hate speech on social media, known as the *Loi Avia* (Boring, 2020). The *Loi Avia* requires that social media platforms and search engines must remove content containing hate speech within 24 hours after being notified. Although the *Loi Avia* was adopted by the French Parliament, it was overturned by the French Constitutional Court a month later. In 2020, the UN released the *United Nations Strategy and Plan of Action on Hate Speech: detailed guidance on implementation for United Nations field presences* (hereinafter referred to as *Strategy*) and identified state actors as bearing the primary responsibility for addressing and combatting hate speech.

The C.O.N.T.A.C.T. project, an analysis of the comments on newspaper websites, however, suggests that hate speech in CMC is not necessarily intended to elicit antagonism between social groups (Assimakopoulos et al., 2017).[9] It found that hate speech stems from

[9] The C.O.N.T.A.C.T. project (JUST/2014/RRAC/AG) was carried out between 2015–2017. The data was collected in ten EU member states (i.e., Cyprus, Denmark, Greece, Italy, Lithuania, Malta, Poland, Romania, Spain and the UK). C.O.N.T.A.C.T.

stereotypes about certain social groups. *Stereotype* refers to a set of messages and beliefs relating to a certain type of object, in particular a social group, elaborated into a distinctive, coherent and stable image (Assimakopoulos et al., 2017). Prejudice is then formed based on the stereotype, when people create hostile and harmful personal opinions about a social group (Assimakopoulos et al., 2017). Since people fear that their opponents may threaten their own self-schemata and expose their weaknesses as stereotypes and prejudices, they tend to desperately reject anything that is opposed to their opinions, to place a high value on the characteristics of their own social groups (e.g., intelligence, civilisation, historical origins) and to vilify/deny those of their opponents. Hate speech is thus used to defend people and their social group against their opponents, rather than fuelling social hatred (Assimakopoulos et al., 2017).

1.3.2 Responses from data scientists

Data scientists focus on automating the detection of conflictual content in virtual contexts (e.g., hate speech, cyberbullying, personal attacks) by developing and optimising computational detection models (Salminen et al., 2020). They regard messages as a collection of "a set of terms and their frequency" (Pereira-Kohatsu et al., 2019, p. 5) and commonly employ lexicon-, syntax- and semantics-based classifiers or features to identify the textual patterns in the message (ElSherief et al., 2018; Pereira-Kohatsu et al., 2019; Salminen et al., 2020; e.g., Araque et al., 2017; Davidson et al., 2017; Saleem et al., 2017; Mondal et al., 2017; Waseem et al., 2018; Quijano-Sánchez et al., 2018; Sahin et al., 2018; Sahlgren et al., 2018; Salminen et al., 2018ab). Lexicon-based classifiers, also known as keyword-based classifiers, function to "check the occurrence of words in the revised text" (Pereira-Kohatsu et al., 2019, p. 5) by using a predefined dictionary. Syntax-based classifiers employ more sophisticated feature representations based on syntactic knowledge such as BOW, N-grams or POS (Pereira-Kohatsu et al., 2019). Unlike the first two classifiers, semantics-based classifiers preserve the semantic relationships between linguistic terms by creating embeddings of a word. That is, they label "the size of the neighbourhood of a word, i.e., the other words that surround it and that define its context" (Pereira-Kohatsu et al., 2019, p. 5). However, these text-based detection models cannot effectively distinguish between expressions which are offensive and those which are not because they often detect

stands for "creating an On-line Network, monitoring Team and phone App to counter hate crime tactics". For more information about the C.O.N.T.A.C.T. project, visit the C.O.N.T.A.C.T. website at: http://reportinghate.eu/en/.

non-offensive expressions as conflictual and offensive ones as harmonious due to the presence of certain words or phrases (e.g., Rajadesingan et al., 2015; Nobata et al., 2016; Watanabe et al., 2018; Salminen et al., 2020).

Expressions that may be perceived as conflictual often take various forms (e.g., Rajadesingan et al., 2015; Nobata et al., 2016; Watanabe et al., 2018; Salminen et al., 2020). There may be a textual feature (e.g., *fuck*, *shit*, *bitch*, *asshole*) or a multimodal feature (e.g., 🤮, 👎, 🖕, 🤐) and textual features may not strictly follow the orthography. For example, the terms *FUCK*, *f****k*, *fucking*, *FUUUUUUUUUCK* and *F* are all variables of the word *fuck*. Second, what appears to be inharmonious may not be used to provoke conflict and vice versa. For example, although the emoji of '🙂' represents a slightly smiling face, it is also often used to be "patronising, passive-aggressive or ironic" ("Emojipedia—Slightly Smiling Face", 2021).[10] In addition, slang and idioms used to offend often quickly and constantly evolve on the Internet. For example, the word *boomer*, an informal term for the word *baby boomer*, refers to people who were born during the time of baby boom, approximately two decades following World War II ("Dictionary—Boomer", 2022). In 2019, an older man in a TikTok video publicly criticised younger generations for being destroyed by social media and participation trophies. In response, the phrase "OK boomer" was used as an Internet slang to satirise the old man for being anachronistic and closed-minded ("Dictionary—Boomer", 2022).[11] This phrase is currently used to criticise old-fashioned people who resist new ideas and changes (e.g., technological change, climate change) and even implies ageism.

The text-based detection models also fail to interpret the impact of interactions on the meaning of language (e.g., Rajadesingan et al., 2015; Nobata et al., 2016; Watanabe et al., 2018; Salminen et al., 2020). The same word or phrase may contain numerous meanings in divergent contexts (e.g., different social media, different stages of a conflictual interaction with varying conflict intensities, different participants and different accidents that occur in real life). People with distinct virtual identities or who belong to different online communities may consider what is offensive based on very typical standards. Such cognitive differences may even be influenced by people's real-life identities, life circumstances and educational experiences. For example, the phrase 'Go fuck yourself' is usually used to stop someone from being offensive after being seriously offended by them. However, when used

[10] Accessed, 22nd February 2022: https://emojipedia.org/slightly-smiling-face/
[11] Accessed, 22nd February 2022: https://www.dictionary.com/browse/boomer

among friends in certain circles, it may function as an endearing term, wishing the other party a good time or saying goodbye. In some cases, it can also be used as a measurement of an incomprehensible amount such as "I drank a go fuck yourself amount of coffee at work!!!!" (Thatfunkymotherfunker, 2010).[12] Based on the context at hand, people select the most appropriate forms of language that accurately express their perspectives, attitudes, desires, feelings, emotions, intentions, etc. This suggests that whether a message is meant to cause conflict or whether it actually serves to intensify the conflict may not necessarily relate to its form or the linguistic features it contains.

In order to optimise the automatic detection of highly contextualised content, sentiment analysis has also been widely adopted. It is used to "determine how an individual or [a] group of people feels and reacts to a specific situation or [a] comment" (Pereira-Kohatsu et al., 2019, p. 4). By detecting and labelling the sentiment polarity of language in social media, data scientists determine users' attitudes. Based on the assumption that words or phrases usually have the same sentiment polarity, regardless of their contexts or implications, this approach still focuses on existing positive/negative words or phrases and heavily depends on their literal meaning (Watanabe et al., 2018). In conflictual interactions, however, people do not usually consider a message offensive based on what it appears to show but rather according to its context. For example, while the phrase 'go fuck yourself' is known to be very offensive, when used between friends it may function as a form of humour and even as an expression of intimacy. It could be argued that it remains a technical challenge for data scientists to train computers to interpret the meaning and function of language as humans do.

1.3.3 Responses from Twitter

To address the proliferation of uninhibited language use on Twitter, the Twitter company has developed a procedure for content moderation, which is "a feature that empowers [Twitter] community to report and vote on [content] that they consider to be spam or abusive" ("Twitter—How to report content and assign moderators", 2022).[13] It has established a set of rules that explicitly prohibit linguistic behaviour that publicly elicits hatred or blatantly violates the law, such as hateful, abusive and violent (verbal) behaviour

[12] Accessed, 22nd February 2022: https://www.urbandictionary.com/define.php?term=Go%20fuck%20yourself
[13] Accessed, 21st March 2022: https://help.twitter.com/en/using-twitter/report-content

(e.g., “Twitter—Hateful conduct policy”, 2022; “Twitter—Sensitive media policy”, 2022; “Twitter—Violent threats policy”, 2022). It suggests that Twitter users promptly mark sensitive content by changing settings when tweeting about such content and encourages users (in some special cases of non-registered Twitter users) to report sensitive content that has not been marked (e.g., “Twitter—Hateful conduct policy”, 2022; “Twitter—Sensitive media policy”, 2022; “Twitter—Violent threats policy”, 2022; “Twitter—How to report content and assign moderators”, 2022). Twitter also established the Trust and Security Team that aims to assess, adapt and develop policies in this area in 2008. Once a post and/or an account is reported, Twitter employees “will randomly be selected to decide if the [content] is offensive” (“Twitter – How to report content and assign moderators”, 2022).[14] They need to categorise the reported content as “Abuse, Spam, Looks OK, or Not Sure” based on their own assessments (“Twitter—How to report content and assign moderators”, 2022).[15] According to the majority rule, Twitter then decides how to deal with the reported cases such as labelling violation, deleting posts and temporarily or permanently suspending accounts (e.g., “Twitter—Hateful conduct policy”, 2022; “Twitter—Sensitive media policy”, 2022; “Twitter—Violent threats policy”, 2022; “Twitter—How to report content and assign moderators”, 2022).

In order to protect the freedom of speech of its users, Twitter’s measures and regulations are specifically designed to combat hate speech. There are, however, a lot of instances of verbal abuse, attacks and threats that are not motivated by a socio-historical hatred towards a particular social group but rather caused by the intensification of debates on Twitter. In terms of hate speech, to my knowledge, the Twitter company has not admitted in any official document or public statement that it employed or is employing any automatic detection software to monitor the use of hate speech on its platform. Twitter argues that it only reviews the reported content and, depending on the level of harm, takes appropriate actions. However, its manual detection is severely restricted by factors such as its employees’ workflow, workload, efficiency and experience. According to a survey conducted by the EU in 2019, Twitter’s response to hate speech is not as effective as other online platforms (including Facebook, YouTube, Instagram and Jeuxvideo.com) and is generally below average. The survey shows that Twitter assesses notifications within 24

[14] Accessed, 21st March 2022: https://help.twitter.com/en/using-twitter/report-content
[15] Accessed, 21st March 2022: https://help.twitter.com/en/using-twitter/report-content

hours in 76.6% of the cases, removes 35.9% of the content that has been notified to Twitter and replies to 43.8% of the notifications it has received; these results are well below the average of 90.4%, 71% and 67.1%, respectively (EU, 2019). In addition, the measures taken by Twitter hardly seriously threaten or damage the actions of users who commit such offences; users who violate Twitter's rules are neither required to pay a fine nor subject to any civil or criminal prosecution from the Twitter company. Their reputations in real life are much less likely to be damaged by being suspended as a user or by having their posts removed.

1.3.4 Responses from linguists

Applied linguists, who were among the first to enter this field, commonly take conflict as an established linguistic phenomenon lacking in movement or change (e.g., Sahlane, 2019; Davies, 2019; Nahajec, 2019; Alaghbary, 2019; Blitvich, 2019; Khan, 2019; De Bres & Franziskus, 2019; Archer et al., 2019; Coinnigh et al., 2019). They aim to investigate linguistic features of conflict and the relationship between the features and the contexts of their occurrence (e.g., social problems, power imbalance, virtual identities). The nature of conflict as a social process that evolves with participants' arguments, however, has been overlooked.

Already in the 1990s, linguists with a pragmatic focus pointed out the problem that in conflict studies, interactions and conflict were typically being "treated in relative isolation" (Briggs, 1996, p. 3; see also Grimshaw, 1990). Thirty years later, the discursive nature of conflict has still not received much attention from linguists, especially those who are interested in pragmatic issues. The second section, entitled 'Interaction in conflict', in *The Routledge Handbook of Language in Conflict* reintroduces the issue of conflictual interactions after this thirty-year hiatus. The introduction of this section, written by Jim O'Driscoll in 2019, is one of the few and also the most recent paper that systematically reviews linguistic research on conflictual interactions. O'Driscoll (2019) surveyed linguistic studies on conflictual interactions in face-to-face interactions and those between acquaintances and between strangers in CMC. He summarised typical research foci and linguistic theories/methods used in linguistic research of conflictual interactions and provided the earliest systematic attempt to establish a theoretical framework and methodological support for studying the interactive nature of conflict.

The studies cited by O'Driscoll (2019) pay great attention to how speakers oppose their opponents by using linguistic strategies, namely the manifestation of opposition.[16] However, the review of linguistic research on conflictual interactions reveals another type of study that treats such interactions as a process of initiating, developing and managing opposition and focuses on their overall structures (which will be discussed in detail in Chapter 2). O'Driscoll (2019) did not consider these studies and hardly investigated how linguistic strategies, through mutual facilitation and constraint, generate conflictual interactions. This slightly skews his summary of the foci, theories and methods of the current studies on conflictual interactions. In addition, O'Driscoll (2019) just outlined these studies briefly, without explaining the intrinsic connections between different lines of research on conflictual interactions. For example, what are the differences and the similarities between disagreements, (im)politeness and conflict? What leads to the different lines of research? Are these lines absolute opposite or related under certain conditions/contexts? O'Driscoll (2019) also failed to propose a thorough theoretical framework or systematic methodological guidance for studying the interactional mechanisms of conflict. For example, how can existing research be used to reveal the developmental mechanisms of conflictual interactions? What theories can account for their dynamic nature? What kind of data is most appropriate for analysing such interactions as a whole? How can such data be collected from social networks and/or social media?

Overall, computer and network technologies have greatly changed human communication. They enable synchronous conversations over long distances, enrich the ways of communicating in daily life and expand the range of people to interact with by creating virtual spaces for free conversations between strangers. However, technology also increases the frequency of and expands the scope of uninhibited linguistic behaviour in CMC, challenges verbal hygiene conventions that are usually observed in face-to-face conversations, changes the standard of appropriate language use and interactional modes and generates and exacerbates a variety of social problems. Therefore, research into the interactional mechanisms of online conflict is able to provide insights into the relative importance of uninhibited versus inhibited linguistic behaviour, thus facilitating the further

[16] To distinguish users who issued the turns this study is analysing and the other users who participate in the interaction, the former are referred to as *speakers* or *the (current) speaker* in this study. Users who respond to the turns under discussion are termed as *recipients* or *the (next) speaker*. Users who are responded to by the turns under discussion are named as *the (previous) speaker*. It should also be noted that on Twitter the recipient may be the same individual as the previous speaker; if other users respond to the current speaker, the recipient may also be a different person.

development of pragmatic research and theories building on inharmonious linguistic phenomena. By taking into account technology as a factor influencing human communication, this study also reveals how technology affects peoples' ideologies of language through constraining and shaping distinctive ways of speaking and communicating on the Internet platforms. It thus contributes to integrating a posthumanist and sociolinguistic approach to the research framework in pragmatics and facilitates the development of pragmatic research towards a more social and multidisciplinary orientation. By doing so, from a practical perspective, this study raises public awareness about their/others' online linguistic behaviour and seeks to further optimise the detection of inharmonious language use that threatens cyber security and social stability. Finally, the remainder of this chapter provides an overview of this thesis (see Section 1.4) and outlines its overall structure (see Section 1.5).

1.4 Overview of the thesis

Section 1.3 suggests that there are two major issues that need to be addressed in this study. On the one hand, conflict is often characterised by interactivity and dynamicity because it emerges as communication begins and develops and as participants' debates progress. At the end of a conversation, even though the deep-rooted conflict between participants may remain irreconcilable, the conflict manifested by language does terminate. This discursive nature of such protracted conflict and the specific use of inharmonious language in such conflictual interactions have not been sufficiently examined in interactional studies so far. This absence has also indirectly led to the ineffectiveness of current computational models in detecting conflictual content as they still rely heavily on linguistic features to train classifiers. On the other hand, network technologies create a virtual space for people to interact with others across time zones and geographical locations and challenge many classic perceptions of natural language and human communication, including definition, categorisation, rules, etc. They also obscure people's sensitivity to and scruples about the real-life constraints imposed by social norms, administrative regulations and laws. This results in a large amount of uninhibited linguistic behaviour being easily observable and documented and further establishes the stereotype that CMC is discourteous. However, this difference generated by technology between face-to-face and online interactions has not been adequately investigated in current interactional studies. Researchers still tend to treat the Internet and conflict as separate social prerequisites for

human communication, and the role of technology in shaping human communication and language use continues to be overlooked.

To address these issues, the present study takes an interactive perspective on conflictual interactions in CMC, viewing protracted online conflict as an interactive process. It assumes that every linguistic element in online conflictual interactions is constantly evolving because it is embedded in a dynamic linguistic context. It aims to explore the developmental mechanisms of conflictual interactions in CMC. To fulfil this aim, this study attempts to deconstruct the dynamic interactional process by proposing several interactional principles to account for how CMC users maintain the continuous progression of conflictual interactions and how they adjust the developmental tendencies of the interaction. On the one hand, this study conducts a set of meticulous observations and analyses of the features and structures of conflictual interactions in one interactive context: Twitter. It identifies the textual and multimodal features employed by Twitter users to engage in conflictual interactions and the purposes for using these features. It also explores the stages in conflictual interactions on Twitter, their characteristics and their functions. On the other hand, this study investigates the interplay between linguistic features and interactional structure. That is, it examines how Twitter users employ linguistic strategies to shape the structure of conflictual interactions and how ongoing conflictual interactions affect users' choice and design of linguistic strategies. The study argues that linguistic strategies should not be excluded from the analysis of the overall structure of conflictual interactions and the role of strategies in shaping the structure should not be overlooked. This perspective constitutes the central idea of this study. If conflictual interactions are taken as a bustling 'highway', their linguistic strategies and developmental mechanisms serve as various 'vehicles' that move along the highway and the rules that guide these vehicles on the highway, respectively. This study is concerned with how Twitter users drive their vehicles, called linguistic strategies, safely and smoothly on the highway, called conflictual interactions, by following traffic rules. It is also interested in the approaches that the users employ to respond to 'unexpected situations' encountered on the highway (e.g., sudden stops, reversals of direction, speeding) and the reasons for these situations.

In order to investigate the workings of conflictual Twitter interactions, this study therefore needs to answer the following two questions: (1) how do people use linguistic strategies to initiate and progress conflictual interactions on Twitter? (2) How do people

engage in conflictual Twitter interactions under the constraints and reshaping of Twitter algorithms? To address these two issues, this study collected two sets of data: conflictual interactions on Twitter and the results from questionnaire surveys. It applied corpus linguistics and built a corpus of conflictual interactions on Twitter, entitled Conflictual Twitter Interactions Corpus, in a coherent and systemic manner. Two surveys were also conducted to establish a reasonable annotation standard of the corpus and to obtain additional information about the reasons, attitudes and impact of language use in conflictual Twitter interactions, thus minimising the impact of researcher bias on analysis. In terms of the corpus data, this study employs discourse analysis (DA) and Brown and Levinson's (1987/1978) facework as its analytical methods. It first identifies the linguistic strategies employed by Twitter users to do conflictual interactions and preliminarily examines their linguistic nature. It then provides an in-depth analysis of the role of these strategies in driving conflictual interactions based on a case study. Specifically, it examines the reasons why strategies are selected by Twitter users and the impact of their strategies on the design of the following turn. Finally, the study investigates the most commonly adopted makeup of linguistic politeness and its function, especially how it intensifies or mitigates the conflict.

1.5 Structure of the thesis

The present study consists of eight chapters and is organised as follows: Chapters 2 and 3 provide an overview of existing literature on conflictual interactions. They perform two functions: the establishment of a theoretical basis and methodological framework for studying the dynamic nature of online conflictual interactions. Chapter 2 critically examines existing conceptualisations and theories of conflictual interactions and reviews research on different types of conflictual interactions (including those that occur in face-to-face interactions and in CMC). It begins with a survey of strategies that speakers employ to oppose their opponents, including by refuting the opponents' opinions (i.e., disagreements) and/or by attacking the opponents (i.e., personal attacks). It then focuses on the dynamic unfolding process of conflictual interactions and, in particular, the roles that linguistic strategies play in starting and terminating interactions. The review shows that there is currently much less work on online conflictual interactions than on offline conflictual interactions and that there is even less research that examines the interconnections between the nature of linguistic strategies and the structure of conflictual interactions. Based on these gaps in the literature, three specific tasks that need to be addressed in this

study are posed in Section 2.3 in Chapter 2: (1) identifying linguistic strategies for (de-)escalating conflictual Twitter interactions, (2) exploring the role of these strategies in the progress of conflictual Twitter interactions and (3) examining the common makeup of linguistic politeness and its discursive functions in conflictual Twitter interactions. The discussion also suggests two sets of data that have to be collected: conflictual Twitter interactions and follow-up surveys, and it reveals an analytical framework for studying the developmental mechanisms of conflictual interactions, that is to combine discourse analysis and (im)politeness theories.

Chapter 3 continues to survey these methodological approaches. First, it establishes corpus linguistics as theoretical support for collecting interactive data and for building a corpus. It discusses the issues and possible solutions for building a corpus of conflictual Twitter interactions that is amenable to a qualitative and a quantitative analysis. Second, it selects DA, combined with Brown and Levinson's (1987/1978) facework as appropriate approaches for analysing the corpus. It critically discusses the strengths and weaknesses of each approach and comprehensively compares the two approaches with *conversation analysis* (hereinafter referred to as *CA*) and impoliteness theories, respectively. They are each suitable for exploring speakers' use and representation of linguistic strategies, for investigating the structure of conflictual interactions and for examining the shaping of the Twitter algorithms on users' language use and interactional patterns. However, they share the problem of overly fragmenting adjacency pairs (APs) from the interactions as a whole. Combining the two becomes the most appropriate approach for exploring the developmental mechanisms of conflictual Twitter interactions. That is, to concretely examine the impact of the current speaker's interpretation of the previous turn on their use of strategies and to examine the impact of the current speaker's strategies on the design and presentation of the next turn. Third, this chapter suggests that questionnaire surveys are suitable for verifying researchers' identification of linguistic strategies and for providing additional information such as Twitter users' perceptions and attitudes to conflictual interactions to complement the textual analysis.

Chapter 4 discusses the methodology of this study in more detail and presents the data, as well as the data collection and analysis methods used for investigating the features and structure of conflictual Twitter interactions. It addresses four issues: first, based on prior studies, it deepens the analysis of how Twitter interactions work, exploring their structures

and key features. Second, it introduces a Twitter data classification criterion to capture a representative range of Twitter interactions. This standard helps to ensure that the conflictual Twitter interactions in the corpus do not merely originate from highly controversial topics or Twitter users. The purpose is to simulate as closely as possible the realistic occurrence of conflictual interactions on the Twitter website. Third, it details the procedures for building a corpus of conflictual Twitter interactions, including: (1) the steps of collecting targeted tweets, (2) the steps of expanding targeted tweets into Twitter interactions, (3) the steps of distinguishing conflictual exchanges from harmonious ones and (4) the steps of annotating the linguistic strategies used by Twitter users to do conflictual interactions, including the design and results of a survey used to validate the annotation principle. Fourth, it explicitly illustrates how to explore the developmental mechanisms of conflictual Twitter interactions by using DA and Brown and Levinson's (1987/1978) facework. It also discusses the design of another survey that functions to obtain Twitter users' attitudes towards language use in conflictual interactions and explains how to use the survey to supplement the analysis of linguistic politeness based on Brown and Levinson's (1987/1978) theory.

Chapters 5, 6 and 7 present the linguistic analysis, addressing the three issues that emerge as salient in the literature review (see Chapter 2): (1) the identification of (de-)escalation strategies, (2) the investigation of the working mechanisms of the strategies and (3) the exploration of the primary makeup and modifying functions of linguistic politeness. Chapter 5 initially identifies escalation strategies in conflictual interactions, including disagreements and personal attacks. It then discusses the reasons for the absence of de-escalation strategies in the corpus. A preliminary investigation into the discursive nature of the escalation strategies is also provided. This chapter reveals the cooperative principle that Twitter users employ to engage in conflictual interactions: they successively switch their role from being a recipient of the (previous) turn to being a speaker of the next turn. The turn constructions of the speaker and the recipient are given significance in revealing the speaker's discursive intent and in revealing the functions of the speaker's turn in conflictual interactions, respectively. The key role of the recipient's turn design in response to the (previous) speaker's strategic use and strategic goals is thus initially proposed. This sets the scene for the following chapter. This chapter also highlights one

stereotype-breaking fact: personal attacks perceived by the recipient appear to drive rather than to hinder the development of conflictual interactions.

Given the lack of linguistic studies on personal attacks, Chapter 6 provides an in-depth analysis of the nature of personal attacks perceived by the recipient and their role in generating conflictual Twitter interactions. The purpose is to further explore the function of linguistic strategies in conflictual interactions. This chapter deals with the following issues: first, it analyses the meaning of personal attacks, by quantifying the occurrence of the posts perceived as a personal attack in the corpus. Second, it provides insights into these personal attacks, particularly their distribution, their linguistic nature and the factors that determine their occurrence. Thirdly, it concludes with an in-depth analysis of the ways in which these personal attacks drive conflictual Twitter interactions, including their impact on the recipient's response and on the nature of conflictual interactions as a whole.

Chapter 7 explores the role of linguistic politeness in conflictual interactions. It begins by defining linguistic politeness in conflictual Twitter interactions as a modifier that functions to adjust (i.e., intensify or mitigate) the intensity of the speaker's conflict with the recipient. It then provides an in-depth analysis of various types of linguistic politeness methods employed by the speaker when doing conflict on Twitter, based on Brown and Levinson's (1987/1978) facework. Finally, it compares in detail linguistic politeness in conflictual Twitter interactions and in face-to-face interactions and discusses the implications of this difference for conflictual Twitter interactions.

Chapter 8 concludes the present study. It first summarises the findings and answers the research question about how conflictual interactions on Twitter are progressed by the language use of their participants. It then discusses the implications of these findings in advancing pragmatic research, in improving the recognition of the impact of technology on natural language and human communication, in informing the establishment of methodological frameworks for future CMC studies and in detecting and controlling uninhibited linguistic behaviour on the Internet. A final comment is provided on the constraints of this study and prospects for future research.

Chapter 2 Reviewing Conflict Studies

2.0 Introduction

Chapter 1 argued that the discursive nature of conflict was first noticed by linguists around thirty years ago, see Grimshaw (1990), Briggs (1996) and Pearce and Littlejohn (1997). However, due to the Anglocentric orientation of research on interactions that takes harmony as a norm in interpersonal interactions (Levisen, 2019), conflictual interactions have been marginalised in interactional studies until quite recently. With the rise of interest in conflictual interactions due to their prevalence on the Internet, a growing number of researchers, including legal professionals, IT practitioners, data scientists and applied linguists, have since acknowledged the discursive nature of conflict and have begun to study the relation between social conflict (including both online and offline conflict) and language use (e.g., RADAR, 2015; EU, 2016, 2019; Mondal et al., 2017; Sahlgren et al., 2018; Sahlane, 2019; Alaghbary, 2019; Blitvich, 2019; Khan, 2019; De Bres & Franziskus, 2019; Salminen et al., 2020; UN, 2020). They have also been devising laws and regulations against linguistic behaviour that serves to incite social hatred, developing feature-based classifiers for automatic online conflict detection and analysing social issues and power imbalances in conflictual interactions. However, they still regard conflict in their research as a static and pre-existing social state that is realised by linguistic features whose aim is to contaminate verbal hygiene. So far, the impact of network technologies on human communication and the interactive nature of conflict have not been extensively investigated. This study aims to contribute towards filling these two gaps by exploring the linguistic mechanisms that trigger the occurrence and development of conflictual interactions on Twitter under the constraints and reshaping of Twitter algorithms.

This chapter surveys linguistic research on conflictual interactions in both face-to-face and online contexts. The aim is to preliminarily explore how to study the developmental mechanisms of conflictual interactions on Twitter, including identifying the main research issues, open questions and suitable methodological approaches. This chapter broadly divides linguistic research on conflictual interactions into two broad categories: research on conflictual linguistic strategies and research on the structure of conflictual interactions.

Section 2.1 discusses linguistic research on the linguistic strategies that people employ when engaging in conflictual interactions. The discussion focuses on two major types

of conflictual linguistic strategies: disagreements (i.e., speakers counter their opponents' opinions; see Section 2.1.1) and personal attacks (i.e., speakers belittle their opponents or deprive the opponents of their rights; see Section 2.1.2). It is also concerned with how speakers express these strategies through the use of linguistic (im)politeness and how contextual factors (e.g., cultural, social and demographic factors) impact speakers' use of strategies and how their strategies influence the relationship between interlocutors. Section 2.2 reviews linguistic research on the structure of conflictual interactions from a holistic view. It regards such interactions as a dynamic process and investigates how these interactions unfold in a chronological order, especially how conflict is initiated and terminated by speakers' linguistic strategies.

2.1 Focus 1: conflictual linguistic strategies

This section reviews thoroughly linguistic studies that examine how speakers use strategies to defend themselves and to challenge their opponents' viewpoints in interactions. They may be broadly divided into two categories, including: (1) studies that focus on how speakers negate others' points of view (see Section 2.1.1) and (2) studies that deal with how speakers attack others (see Section 2.1.2). This review shows that the studies surveyed by O'Driscoll (2019) pursue this focus (see Section 1.3.4 in Chapter 1). It also suggests that disagreement studies, especially those based on offline data, constitute the majority of linguistic research on conflictual interactions. Only a few studies have examined a particular type of personal attack: hate speech. Each type will be discussed below.

2.1.1 Objecting to the recipient's opinion

Disagreement represents a type of linguistic strategy that is the opposite of agreements (Muntigl & Turnbull, 1998; Rees-Miller, 2000; Angouri & Locher, 2012). In some studies, it is also referred to as arguments (e.g., Schiffrin, 1985; Schegloff, 2007; Al-Hindawi et al., 2018), debates (e.g., Johnson & Johnson, 1985; Yuan et al., 2019; Moernaut et al., 2020), disputes (e.g., Nelson, 2001; Williams, 2005; Jeong, 2008; Tracy & Hodge, 2019), confrontations (e.g., Brown, 1990; Hutchby, 1992), opposition (e.g., Kakavá, 2002) and agonism (e.g., Tannen, 2002). Although issues such as how these strategies differ in definition and function as well as their relation to (im)politeness have been subject to considerable debate, it is generally accepted that they cannot be easily fully distinguished from each other. The present study acknowledges that there are subtle differences between

these strategies. It is, however, not the purpose of this study to neatly distinguish these terms. This study focuses on developing a generalisable interactional model from the data, which requires that researchers should refrain from focusing excessively on the differences between terms, and thus adopts the most commonly used term, 'disagreement'. Disagreements are employed when a communication process is disturbed by the divergence in interlocutors' beliefs about a particular issue (Stapler, 1995). They function to indicate that the current speaker takes a position that is opposed to that of the previous speaker (Stapler, 1995; Rees-Miller, 2000; Locher, 2004; Angouri & Locher 2012; Sifianou, 2019). However, there are cases where one party does not consider the other side to be their opponent, which contrasts with the other party's perception. Stapler (2015) found that sometimes, one of them believes that a consensus cannot be reached while the other assumes that they share common ground. In short, disagreements are used by the current speaker, when they misunderstand or dissent from the claim uttered or presumed by the previous speaker, to communicate a negative implication of the previous speaker's assertion.

According to Angouri and Locher (2012), there are three main foci in disagreement studies. The first one aims to identify the different ways of expressing disagreements, such as what specific linguistic features speakers employ, whether their disagreements go bald-on-record or off-record and whether they perform redressive actions in their objections (e.g., Schiffrin, 1985; Muntigl & Turnbull, 1998; Scott, 2002; Locher, 2004; Shum & Lee, 2013; Black & Wiederhold, 2014; Ben-Menachem & Livnat, 2018; Almutairi, 2021). Linguists commonly carry out the identification based on (im)politeness theories, particularly Brown and Levinson's (1987/1978) facework, regarding face as "basic wants, which every member knows every other member desires and which in general is in the interests of every member to partially satisfy" (Brown & Levinson, 1987/1978, p. 62). They argue that objecting to the opponent's viewpoint functions as a *face-threatening act* (FTA) as this would cause a threat to the opponent's face but this face-damage can be redressed by using linguistic politeness strategies. The second focus examines how social factors such as gender, age, cultural background, social class and interactional goals influence speakers' specific ways of negating others' viewpoints (Angouri & Locher, 2012; see also Kakavá, 2002; e.g., Goodwin & Goodwin, 1987; Johnstone, 1989; Brown, 1990; Herring, 1994, 1996b; Kakavá, 2002; Edstrom, 2004; Williams, 2005; Habib, 2008; Behnam & Niroomand 2011; Koczogh, 2012;

Yang, 2013; Decock & Spiessens, 2017a; Suroiya, 2017; Bova & Arcidiacono, 2018; Toomaneejinda & Harding, 2018; Fernández-Amaya, 2019; Khammari, 2021). By employing DA, researchers insist that people use language to present different perspectives and understandings to express their sense of belonging to different social groups and to integrate into the social world. The last type, usually embedded in the previous two foci and methods, explores the outcomes of different disagreement expressions and their impacts on the relationship between interlocutors (e.g., Schiffrin, 1984; Tannen & Kakavá, 1992; Georgakopoulou, 2001; Tannen, 2002; Locher, 2004; Habib, 2008; Fernandez-Amaya, 2019). Researchers are particularly interested in whether speakers' disagreements aggravate or mitigate the potential threat to their opponents' faces and whether they make interlocutors intimate or apart.

Most disagreement studies are based on conflictual interactions in face-to-face contexts (e.g., Schiffrin, 1985; Goodwin & Goodwin, 1987; Johnstone, 1989; Muntigl & Turnbull, 1998; Brown, 1990; Rees-Miller, 2000; Georgakopoulou, 2001; Kakavá, 2002; Ben-Menachem & Livnat, 2018). For example, Ben-Menachem and Livnat (2018) studied Havruta conversations among Jewish women. Havruta conversations refer to a linguistic activity that arises in the study of the Talmud, an ancient Jewish religious text, with the goal of maximising mutual comprehension between participants. They are characterised by a highly confrontational and argumentative conversational pattern and offer a high level of tolerance of disagreements. Since Havruta conversations were traditionally conducted by men, Ben-Menachem and Livnat (2018) aimed to identify, for the first time, the disagreement strategies that Jewish women employ in these interactions. They observed 21 Havruta conversations and analysed how one party opposes the other in detail by using Brown and Levinson's (1987/1978) face-saving theory combined with DA. In order to ensure that their findings are not too subjective and biased, they also did a follow-up interview with each participant asking them to reflect on their language use in the conversation. The result shows that Havruta conversations involving women are much less argumentative than those involving men. Only when female speakers consider that they possess a comparable level of Talmudic knowledge as the recipient, do they employ disagreements frequently. They also generally mitigate disagreements by explaining reasons, inserting personal-plural verbs (e.g., nikra "read(pl.)") or hedges (e.g., lefi da'ati "in my opinion") and using agreement markers (e.g., sababa "great", nakhon "right"). There is only one conversation in which the recipient

reported that their face was hurt by the speaker because the speaker opposed the recipient bald-on-record.

Just as Ben-Menachem and Livnat's (2018) study shows that offline disagreement studies tend to include two types of data: a small number of offline conflictual interactions that occur in particular scenarios and follow-up interactions (e.g., interviews, questionnaires) with the participants from whom the interactions are collected. DA and (im)politeness theories are usually combined to analyse the former data. A follow-up survey is conducted to independently verify the findings of the textual analysis and to gather auxiliary information. The combination of these approaches has also been widely employed in disagreement studies that are based on interactional data in virtual contexts (e.g., Shum & Lee, 2013; Decock & Spiessens, 2017b; Fernández-Amaya, 2019; Almutairi, 2021). These studies tend to make some minor adjustments to data collection and analysis to take account of the contingencies of virtual contexts. Shum and Lee's (2013) work is a case in point.

Shum and Lee's (2013) study is one of the first to examine disagreements in online conflictual interactions. Its interactive data comes from Hong Kong social media, which is highly similar to conflictual Twitter interactions. It employs almost every widely used method of data collection and analysis in online disagreement research. Shum and Lee (2013) identified the scope of the data in terms of context, topic and participant based on their research aim. To explore how people generally disagree with others in virtual contexts, they selected two popular online fora in Hong Kong and four posts that sparked heated discussions. These posts contain a controversial topic involving some moral judgements and values and also a general social topic. Unlike Ben-Menachem and Livnat (2018), Shum and Lee (2013) regarded all users who responded to these four posts as participants and did not further restrict the participants' demographic features. As a result, a total of 317 replies were collected. Their data analysis employed a similar approach to Ben-Menachem and Livnat (2018). They carried out a fine-grained and detailed analysis of the collected interactions to identify disagreement strategies and conducted a follow-up questionnaire survey. Due to the anonymous nature of the Internet discussion fora, they randomly recruited 30 participants for the survey, rather than contacting users who had engaged in the interactions they collected. Fifteen of them also participated in the interviews. The aim of this survey was to confirm researchers' identification of disagreement strategies and to

understand how users interpret these strategies in virtual environments. Shum and Lee (2013) identified 11 types of disagreement strategies, which are displayed in Table 1, and established the frequency of each type in the conflictual interactions under investigation. They also calculated the mean score of participants' ratings of each type on three dimensions: (dis)politeness, (dis)appropriateness and negative/positive. They noticed that in these Hong Kong Internet discussion fora, speakers tend to state opinions directly without mitigation. In response to unacceptable opinions, speakers are likely to behave bald-on-record, rebuking interlocutors, giving negative comments and opinions and even using curses, particularly when discussing controversial topics. This language use was tolerated in the case where speakers' objections do not violate moral standards, a set of universal rules used by members of the same social group to distinguish between what is proper and what is improper.

Table 1

Disagreement strategies identified by Shum & Lee (2013)

	Strategy	Explanation & Instances
1	Giving negative comments	The current speaker makes comments on the previous speaker's opinion using a comparatively negative tone.
2	Using short vulgar phrases	The current speaker uses short vulgar phrases such as taboo words, swear words, abusive or profane language.
3	Raising rhetorical questions	The current speaker adds negative tag questions or questions with negative interrogatives.
4	Making a personal stance	The current speaker denies association or common ground with the previous speaker.
5	Making an ironic statement	The current speaker makes ironic statements.
6	Cursing	The current speaker warns, threatens or tries to frighten the previous speaker by predicting that a certain consequence or detrimental event will occur to the previous speaker.
7	Giving opposite opinions	The current speaker gives an opinion that is contrary to the previous speaker's claim.
8	Rewording	The current speaker makes minor changes to or reconstructs the previous speaker's original claim.
9	Giving personal experience	The current speaker refers to their previous experience to substantiate a disagreement with the previous speaker.
10	Giving facts	The current speaker quotes external information such as quotations, statistical information, pictures and videos.
11	Reprimanding	The current speaker tells the previous speaker that their action/behaviour or attitude is not approved and the message may involve emotion.

In terms of Twitter, it serves as a social media platform ("Twitter—Twitter privacy policy", 2022) and provides a virtual space where people can escape from the constraints of off-line life such as social status and personal identities, hide their real personal information or even create virtual identities and freely express their views, attitudes and preferences

(Ilbury, 2020). Therefore, Shum and Lee's (2013) methods of data collection and analysis are highly relevant to a study on Twitter. To study conflict on Twitter, researchers may collect two types of data: conflictual Twitter interactions and follow-up surveys (e.g., questionnaires, interviews). To guarantee the representativeness of the interactions under study, they can select both highly and less controversial topics to avoid demographically and socially restricting participants as much as possible. This approach is suitable for ensuring that findings are relatively general and not limited to certain social groups or settings. The anonymous nature of Twitter also makes it possible for researchers to survey random Twitter users, not necessarily those who participated in the collected interactions. This approach helps to obtain Twitter users' evaluations, attitudes and preferences of some salient issues identified or some findings derived from the textual analysis and other information that is not readily accessible in texts. Combining (im)politeness theories with DA constitutes a suitable method for analysing the textual data. DA provides insights into how Twitter's unique algorithms shape users' interactions and language use and how they further influence users' relationships with each other. (Im)politeness theories are useful for exploring the makeup and function of speakers' language use when doing conflictual interactions. For example, what kinds of conflictual linguistic strategies do speakers employ? What linguistic features do speakers use to express these linguistic strategies? Do they manifest these linguistic strategies in an on-record or off-record manner? Do they use linguistic (im)politeness when designing these linguistic strategies?

Shum and Lee (2013), however, collected conflictual interactions from just four posts which appear quite restrictive and did not elaborate on how these posts were selected. Developing a systematic approach and/or theory used for selecting and collecting Twitter interactions is thus required. As for data analysis, Shum and Lee (2013) only examined how speakers reject their opponents' viewpoints but did not explore how speakers attack their opponents, which is also quite common in conflictual interactions. Although there are no statistics that I am aware of that demonstrate that attacks are more frequent on the Internet than in face-to-face contexts, the ubiquitous verbal violence does constitute a stereotype of CMC, particularly of Twitter interactions, rather than of face-to-face contexts. There is usually a high standard of verbal hygiene in face-to-face interactions, notably in English-speaking cultures (see Section 1.1 in Chapter 1). This social norm requires that individuals

are supposed to be kind and friendly when communicating and to heavily frown upon verbal attacks. The next section surveys linguistic research on verbal attacks.

2.1.2 Attacking the recipient

In conflictual interactions, it is quite common for speakers to use profanity gradually or suddenly to insult their opponents as being dirty or as having some evil personalities or to intervene or to restrict their opponents' actions. However, few linguists have given particular attention to this phenomenon to date, in addition to hate speech (see Section 1.3.1 in Chapter 1). Although hate speech is one type of attack, personal attacks arise from more general causes and their target persons are more widespread. According to Rosé (2017), the present study refers to this type of linguistic strategy directed at the opponent as *personal attacks*.

Some researchers examined how speakers decline their opponents' points of view by criticising the opponents' characteristics. For example, the "aggressive expression and insults" strategy in Ben-Menachem and Livnat's (2018, p. 39) study, the use of "short vulgar phrases" strategy in Shum and Lee's (2013, p. 58) study and in Fernández-Amaya's (2019, p. 1072) study. However, their focus remains on disagreements, rather than on personal attacks. For example, Ben-Menachem and Livnat (2018) found that speakers sometimes oppose their opponents by attacking the opponents directly and explained this with an example of Ruti against Nofet, female students studying the *Talmud*. In Excerpt 1, Ruti questions the credibility of Nofet's understanding of the *Talmud* by attacking Nofet's lack of competence. Ben-Menachem and Livnat (2018) argued that in Turn 35, Ruti opposes Nofet by questioning the reliability of Nofet's affirmations in Turns 32 and 34 as these affirmations are not supported by the text of the *Talmud*. Ruti thus casts doubt on Nofet's skills in studying the *Talmud*, "that's like, your interpretation" (Turn 35). In the following interview, Nofet suggests that she took Ruti's Turn 35 as an attack, arguing that Ruti's criticism went far beyond a regular argument of opinion. Ben-Menachem and Livnat (2018) concluded that Ruti's sentence, "that's, like, your interpretation" (Turn 35), poses a direct attack on Nofet. They then displayed Turn 57 alone, explaining that Ruti in Turn 57 attacks Nofet publicly again by explicitly denigrating Nofet's *Talmud* study. Ruti's objections against Nofet's characteristics in Turns 35 and 57 were thus interpreted by Ben-Menachem and Livnat (2018) as attacks. But in terms of the strategic goals, Ruti's objections provide a valid and

reliable basis for concluding that the views of Nofet are not credible. The attacks mentioned by Menachem and Livnat (2018) thus still pertain to the notion of disagreement.

Excerpt 1 (Ben-Menachem & Livnat, 2018, p. 39)

Turn	Speaker	Text
32	Nofet	Yes.
33	Ruti	That's not right.
34	Nofet	Yes!
35	Ruti	That's not right, (0.1) you, that's like, your interpretation. It's not right.
Omit Turns 36–56		
57	Ruti	= I know- I- It's odd that that's what it says I think that you are interpreting it the way you do because you don't know what he wants.

Studies of hate speech represent almost the only linguistic research on personal attacks that I was able to find to date. They focus on how language is employed to express hatred against a particular social group that shares or resembles existing or imagined characteristics (e.g., gender, ethnicity, race, beliefs, religion and stereotypes) (see Section 1.3.1 in Chapter 1). Blitvich (2018) mentioned the significance of hate speech in examining conflict at a conversational level, arguing that linguistic strategies in conflictual interactions range from disagreements to hate speech. Linguistic research on hate speech has been in full swing in recent years as hate speech has become "one of the most frequent methods for spreading divisive and discriminatory messages and ideologies" (UN, 2020, p. 3). However, most of these studies come from the discipline of computational linguistics (e.g., Sood et al., 2012; Dinakar et al., 2012; Chen et al., 2012; Dadvar et al., 2013; Kwok & Wang, 2013; Gitari et al., 2015; Burnap & Williams, 2015; Kansara & Shekokar, 2015; Djuric et al., 2015; Nobata et al., 2016; Mondal et al., 2017; Wulczyn et al., 2017; Davidson et al., 2017; Park & Fung, 2017; Adamczak-Krysztofowicz & Szczepaniak-Kozak, 2017; Assimakopoulos et al., 2017; Qian et al., 2018; Zhang et al., 2018; Salminen et al., 2018ab; ElSherief et al., 2018; Watanabe et al., 2018; Salminen et al., 2020). They aim to explore how to detect hate speech effectively using automatic methods and how to improve the accuracy of computational models for hate speech detection. They are particularly concerned with the textual features of hate speech but few of them have explored the discursive nature of hate speech.

However, hate speech is not sufficient to define all types of attacks in conflictual interactions. For example, in Excerpt 2, Liz does not attack Jon for sharing similar characteristics with a vulnerable social group. Liz's attack on Jon undoubtedly does not

belong to hate speech as it is generally understood, but Liz does attack Jon's intelligence by calling him "stupid" (Line 02).

Excerpt 2 (Agler, 2012, p. 3)

Line	Speaker	Text
01	Jon	I think we should raise taxes in the USA
02	Liz	You sir, are stupid!

Graham's (2019) study on conflictual interactions in the *League of Legends*, a multiplayer online battle arena from Twitch.tv, suggests a similar case. Twitch.tv is an online platform on which broadcast streamers play a variety of video games and chat with their viewers. Graham (2019) analysed a video stream and accompanying live chat with a female player, Raihnbowkidz.[17] In the video stream, Raihnbowkidz wore a low-cut top and placed her camera overhead, highlighting her cleavage in the video window. Based on Graham's (2019) discussion, Insanity113 criticises Raihnbowkidz for being sexually promiscuous by calling her "little slut" (Line 01) and tells her to commit suicide. In response, Raihnbowkidz defends her adoption of the boobie-streamer persona. She argues that it is a normal feature of her dress that functions to gain expertise in the game and that she does not intend to engage in a morally questionable act of showing her body for money. In this instance, Insanity113's attack on Raihnbowkidz is motivated by the belief that *boobs* represent a sign of immorality rather than discrimination against women. This attack is clearly not an instance of hate speech. Graham (2019. p. 319) referred to this type of interaction with an explicitly harmful purpose as a "true conflict", but she did not specify that what shows in Excerpt 3 is a personal attack nor did she analyse this strategy in depth at an interactional level.

Excerpt 3 (Graham, 2019, p. 320)

Line	Speaker	Text
01	Insanity113	Please kill yourself you little slut

In online conflictual interactions, it is also very common for the current speaker, after arguing about an issue for several rounds, to suddenly attack the previous speaker or a social group that shares similar attributes with the previous speaker. Consider a situation where Users A and B are strangers sunbathing on the same bench. When a furry animal abruptly flashes by from the bushes, User A drops out excitedly: "Wow! What a cute puppy!". User B

[17] Stream refers to "digital data (such as audio or video material) that is continuously delivered one packet at a time and is usually intended for immediate processing or playback" (Accessed, 22nd February 2022: https://www.merriam-webster.com/dictionary/stream#h2). Streamers refer to someone who broadcasts a video for audiences to watch on a stream (Accessed, 22nd February 2022: https://www.merriam-webster.com/dictionary/streamer).

responds with “No! That must be a kitten.” Immediately afterwards, User A opposes User B by arguing that “I’m sure it is a dog because its tail is not as long as that of a cat.” A debate ensues between them about whether that animal is a cat or a dog. As the debate heats up, User B suddenly becomes agitated, saying that “What’s wrong with your eyes? It’s definitely a cat! Idiot.” In this situation, User B’s attack on User A is simply the result of an increasingly heated argument that functions as a reinforcing factor and is not motivated by prejudice against a particular group. This is definitely not a case of hate speech as well.

In summary, as can be seen in Table 2, studies on disagreements and personal attacks are particularly concerned with how speakers employ linguistic strategies to engage in conflictual interactions. For example, what types of conflictual linguistic strategies do speakers use? What resources are involved in expressing these linguistic strategies? Do they employ linguistic politeness? The investigation of these issues forms a very important and fundamental part of any exploration of the developmental mechanisms of conflictual interactions in CMC for two reasons. First, it helps to identify what conflictual linguistic strategies are employed to do conflict with the recipient and to examine in detail how speakers express these strategies. Second, it is suitable for exploring how the algorithm of different CMC modes impact their users’ selection and expression of these strategies.

Table 2

A summary of studies on disagreements and personal attacks

<table>
<tr><td colspan="2">Research Focus</td></tr>
<tr><td colspan="2">Conflictual linguistic strategies (especially disagreements) used in conflictual interactions (especially in off-line contexts).</td></tr>
<tr><td colspan="2">Research Aims</td></tr>
<tr><td colspan="2">1. To identify the types of conflictual linguistic strategies;
2. To explore resources involved in expressing linguistic strategies;
3. To investigate the use of linguistic politeness.</td></tr>
<tr><td colspan="2">Methodology</td></tr>
<tr><td>Data Collection</td><td>Data Analysis</td></tr>
<tr><td>1. Interactional data:
(1) Conflictual interactions on highly & less controversial topics;
(2) The amount of data should not be too small;
(3) Participants should not be demographically or socially limited.</td><td>(1) Discourse analysis
(2) (Im)politeness theories</td></tr>
<tr><td colspan="2">2. Follow-up surveys (e.g., questionnaires, interviews)</td></tr>
</table>

As Table 2 shows, the research methods in these studies are also relevant for the present study. In order to explore how Twitter users generally engage in conflictual interactions, the data needs to constitute a fairly realistic reflection of the ranges and types of conflictual interactions on Twitter. It should include not only conflictual Twitter

interactions on highly controversial topics but also those used in the discussion of less controversial topics. The number of replies included in the collected interactions should also not be too small and the users who take part in these interactions should not come from a demographically or socially limited group. DA and (im)politeness theories offer a useful framework for analysing interactions, which enables researchers to identify conflictual linguistic strategies and their manifestations in interactions. They are also useful for focusing on how Twitter algorithms shape users' language use and for comparing the interactional differences between face-to-face and online contexts. While these studies primarily focus on textual analysis, they also verify Twitter users' attitudes and preferences to the salient findings from texts through follow-up surveys, thus ensuring the reliability of the analysis. The survey may also provide additional information to the textual analysis.

However, the analysis of conflictual linguistic strategies is not able to adequately reveal the linguistic mechanisms that progress conflictual Twitter interactions. It only explains how the current speaker opposes the previous speaker and how their response damages the recipient's face but fails to examine the role of the current speaker's linguistic strategies in the progress of conflictual interactions. It is unclear how their strategies influence the next speaker's language use and how they affect changes in conflict intensity, especially how linguistic strategies escalate or de-escalate the interaction. In order to reveal the developmental mechanisms of conflictual Twitter interactions, an in-depth analysis of the role of linguistic strategies in the interactions is also required. Studies on the structure of conflictual interactions are a type of linguistic research that addresses these issues, which are however relatively rare and were not mentioned in O'Driscoll's (2019) review (see Section 1.3.4 in Chapter 1). However, they are indispensable for analysing the dynamic nature of interactions because they regard conflictual interactions as a ditto focus on the dynamics of these interactions and the relationship between the emergence of different phases within the interactions and speakers' language use. The next section reviews research on the structure of conflictual interactions.

2.2 Focus 2: structure of conflictual interactions

Linguistic studies that treat conflictual interactions as a dynamic process argue that these interactions include not only the specific manifestations of opposition but also the coherent developmental processes of opposition from induction to termination (e.g.,

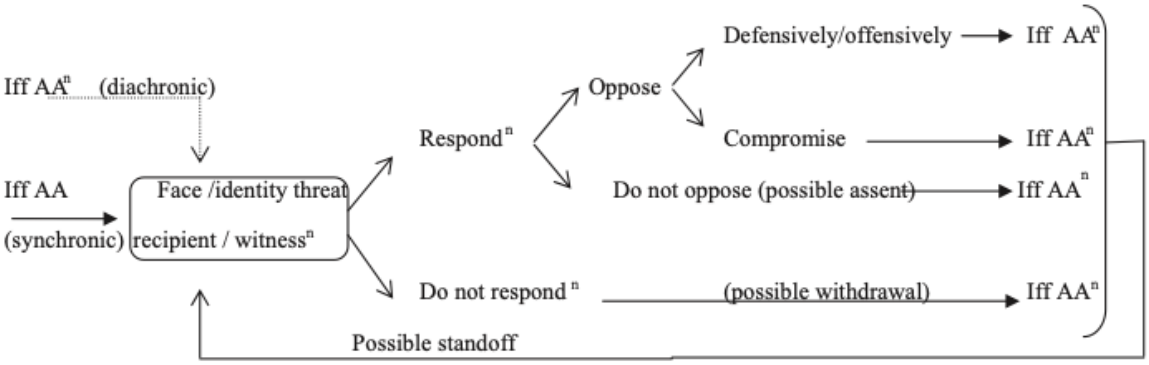
Eisenberg & Garvey, 1981; Vuchinich, 1990; Hutchby, 2001a; Bousfield, 2007, 2008; Nguyen, 2011; Dobs & Blitvich, 2013; Bou-Franch & Blitvich, 2014). They explore how conflictual interactions gradually unfold over time. Just as in studies on linguistic strategies, studies on the overall structure generally use DA as their primary analytical framework but in a rather different way. They do not use DA to explore the interrelation between language use and contextual settings. Instead, they employ DA to discover the structure of conflictual interactions beyond words and sentences. They tend to treat conflictual interactions as a dynamic process consisting of three stages: the beginning of conflict, the end of conflict and the stage in the middle. They aim to explore the key factors that trigger a shift in the nature of interactions from harmonious to conflictual and from conflictual back to harmonious. That is, they investigate linguistic strategies that initiate and terminate conflictual interactions. A case in point is Bou-Franch and Blitvich's (2014) study of conflictual interactions on YouTube. This work proposes the first developmental theory of conflictual interactions in social media between a wide range of participants.

In order to investigate how conflict arises, unfolds and ends in conversations on YouTube, Bou-Franch and Blitvich (2014) built a corpus of conflictual YouTube interactions. Their data are second-hand data obtained from the GenText Corpus. The GenText Corpus is a non-public corpus containing 42 million words created by the University of Valencia and serves to study gender inequality in Spanish and British societies (Bou-Franch & Blitvich, 2014). The data comprises texts extracted from various online media such as digital newspapers and their comments, blog posts and discussions on online fora and YouTube. Bou-Franch and Blitvich (2014) selected one YouTube video in which a group of young members of the LGBT community discusses their sexual preferences, and they collected all the comments in Spanish triggered by that video: 444 in total (about 23,700 words). They noticed that the YouTube algorithms determine the structure of YouTube interactions and these comments were not displayed in the sequence in which they were issued. YouTube users interact by posting new comments or replying to existing comments. Taking the former method, users' comments are displayed at the top of the list of video comments as a direct reply to the video. When responding to an existing comment, users' comments appear below the one to which they are responding and slightly indented to the right. A user may sometimes reply to another comment by posting a new one. As a result, comments on a YouTube video cannot be read in a chronological order. Bou-Franch and Blitvich (2014)

designed their corpus by sorting out the reply relationships between the comments. To analyse the corpus, they built a model based on some popular studies on the structure of conflictual face-to-face interactions. First, they interpreted the emergence of conflict based on Hutchby's (2001a, p. 126) opinion that "any action is in principle open to being treated as an arguable". They assumed that a conflictual interaction is "triggered by the opposition it receives and should be seen as emergent in sequential interaction" (Bou-Franch & Blitvich, 2014, p. 23). Second, they selected Dobs and Blitvich's (2013) theory to explain the phase where conflictual interactions have begun but have not yet ended. Dobs and Blitvich (2013) observed classroom discussions of eighth grade students. They found that no matter how students express an opposing viewpoint, their objections would cause face-damage to others and receive others' objections in response. In this way, students tend to fall into a closed loop of the transition from the actor of conflict to the receiver of the conflict, resulting in possible standoffs in the interaction. In terms of the termination of conflictual interactions, Bou-Franch and Blitvich (2014) chose Vuchinich's theory (1990). Based on a study of American family dinner talks, Vuchinich (1990) found that conflictual interactions may end in the following four ways: (1) the speaker persuades the recipient to give up their objection; (2) the recipient makes a concession, dropping their opposition; (3) the topic of the interaction changes; (4) the recipient withdraws from the interaction by stopping responding to the speaker and (5) a third party, often with higher authority, requests participants to resolve the dispute. However, when using this model to analyse their corpus, Bou-Franch and Blitvich (2014) found that these theories of offline conflictual interactions were not suitable for conflictual interactions on YouTube. As can be seen in Figure 1, They modified these theories based on what they found in their corpus, thus providing the first explanation of conflictual interactions on social media involving a large number of users. Specifically, if an argument is treated as a face-threat by the recipient (i.e., the previous speaker or other recipients), there are four possible responses. They can adopt the withdrawal strategy and cease to respond. They can also respond without opposition, which shows a possible assent. The last two involve expressing opposition defensively/offensively or in a compromising way. If the response is perceived as a face-threat, a possible standoff may occur, leading to a repeat of the whole process.

Figure 1

Bou-Franch & Blitvich's (2014) conflict model



Note: AA refers to arguable action. Reprinted from "Conflict management in massive polylogues: A case study from YouTube," by P. Bou-Franch and P. G. C. Blitvich, 2014, *Journal of Pragmatics*, 73, p. 33.

Bou-Franch and Blitvich's work (2014) and other studies (e.g., Eisenberg & Garvey, 1981; Vuchinich, 1990; Hutchby, 2001a; Bousfield, 2007, 2008; Nguyen, 2011; Dobs & Blitvich, 2013) on the interactional structure are very instructive for the present study on three aspects. First, in studies on linguistic strategies, researchers' understanding of conflictual interactions is fragmented because they focus on the details of speakers' language use. As shown in Table 3, this contrasts with studies on the interactional structure of conflictual interactions that approach such interactions from a holistic perspective. In this perspective, conflictual interactions are not merely a collection of conflictual linguistic strategies but a dynamic interactive process with different developmental phases. This holistic approach is as important for exploring the developmental mechanisms of conflictual Twitter interactions as the nature and makeup of the linguistic structures used in conflictual interactions.

Table 3

A summary of studies on the interactional structure

Research Focus
The dynamicity and structure of conflictual interactions (especially in off-line contexts);
Research Aims
1. To investigate how linguistic strategies initiate conflictual interactions; 2. To investigate how linguistic strategies terminate conflictual interactions.

(To be continued)

Methodology	
Data	Data Analysis
1. Interactional data (1) Bou-Franch & Blitvich (2014): building a corpus of conflictual interactions on YouTube.	1. Discourse analysis: (1) Exploring the role of speakers' conflictual linguistic strategies in driving conflictual interactions; (2) Investigating how conflictual interactions develop through participants' sequential use of linguistic strategies.

Second, Section 2.1.1 argued that the present study, like Shum and Lee's (2013) work, aims to examine how users engage in conflictual interactions in a general way. However, Shum and Lee (2013) did not suggest any suitable method or theory for collecting and processing such data. Bou-Franch and Blitvich's (2014) study provides a possible solution, namely the building of a corpus. *Corpus* refers to a "collection of machine-readable authentic texts which is sampled to be representative of a particular language or language variety" (McEnery et al., 2006, p. 4). *Corpus linguistics* is a branch of linguistics that excels at addressing questions that require a large volume of naturally occurring linguistic data (e.g., McEnery et al., 2006; McEnery & Hardie, 2013). It provides a set of systematic and scientific approaches for identifying the diversity of linguistic data, sampling linguistic data based on research aims and for collecting, processing and analysing linguistic data (e.g., McEnery et al., 2006; McEnery & Hardie, 2013). Third, the use of DA to analyse conflictual interactions in these interactional studies is also suitable for this study. In addition to being used to analyse relationships between language use and its contexts, DA is also appropriate for studying language-in-interaction in detail. It explores the role of speakers' conflictual linguistic strategies in driving conflictual interactions and how these interactions develop through participants' sequential use of conflictual linguistic strategies. This suggests that DA may not only assist in focusing on the unique environment provided by Twitter algorithms. More importantly, DA may also contribute to an in-depth analysis of how speakers, through an elaborate design of their strategies, initiate conflictual interactions, drive these interactions forward and ultimately terminate them.

However, two issues emerge in these studies on the interactional structure. First, they have paid too much attention to the occurrence and termination of conflictual interactions. They investigate the dynamic nature of conflictual interactions at a temporal level and consider conflictual interactions to be dynamic because the states of interactions constantly change over time. They deconstruct conflictual interactions based on the chronological order in which different states emerge. Conflictual interactions are thus

distinguished into three stages: the onset stage, the termination stage and the in-between stage. They aim to explore how speakers' language use propels conflictual interactions into a new stage. Initiating and terminating conflict through the use of linguistic strategies has thus become a main concern of these studies. They, however, do not examine in detail how conflictual interactions are sustained and prolonged by the successive use of linguistic strategies by participants after the beginning and before the termination. For example, how do participants decide the order in which they speak and the sequence in which they use linguistic strategies? How do they determine what specific linguistic strategies they use and how to express their strategies? If the order of these strategies is altered or if the types or linguistic realisations of some of these strategies are altered, how would the conflictual interactions change accordingly? Dobs and Blitvich (2013) and Bou-Franch and Blitvich (2014) shed some light on this. They noticed that a participant's role constantly changes between being an actor of an FTA and a receiver of the previous FTA. They argued that this role shift ensures that conflictual interactions continue. However, they did not examine this finding in detail. They are predominantly concerned with how speakers express their different views through conflictual linguistic strategies and consider this role shift as a closed loop. This contrasts with the case of Twitter interactions where a user who is opposed by the current speaker may not respond to the current speaker. Other users may assume the role of the recipient of the current speaker. That user here is not the direct receiver of face-damage although they may feel that their face has been threatened by the current speaker. Therefore, although studies on the interactional structure appear to examine how conflictual interactions progress through language use, they do not delve into how these interactions progress step-by-step from short to long and eventually to the end as participants use conflictual linguistic strategies in sequence.

Interpreting the dynamicity leads to the three-stage view in linguistic studies on the overall structure and further to the second issue. These studies have oversimplified the changes during the course of a conflictual interaction. Conflictual interactions are dynamic not only in that their states alter over time but more importantly because their intensity changes in a regular pattern. For example, Marcinowski and Ławrynowicz (2020) collected 50 conflictual interactions on Wikipedia with 249 posts and investigated whether the emotional intensity of these interactions increases over time. They found that the average level of the emotional intensity at the end of conflictual interactions is quite a lot higher than that at the

beginning because all of these interactions ended with personal attacks. Xie et al.'s (2013) study on conflictual interactions between students in online classes is another case in point. They found that students' uninhibited linguistic behaviour during free chats before the class was quickly restricted, due to their nervousness at the beginning of the class. As the discussion deepened, conflict emerged and got trapped in a cycle of reinforcement and suppression because uninhibited linguistic behaviour occurred more frequently. When these actions became uncontrollable, conflict intensified and reached a climax. Finally, since conflict severely disrupted the normal discussion, it was controlled by the third party (i.e., the instructors) and entered the stage of adjournment. This suggests that when a conflictual interaction just emerges, its intensity may be relatively weak because participants only express their different viewpoints while being sensitive to the other party's face-needs. As the conflict between the participants becomes more acute, they may use rude and offensive language or attack the other party, leading to the escalation of this conflictual interaction. When the participants use personal attacks frequently, the intensity of this conflictual interaction may reach a peak. Its intensity may then plummet to zero because it comes to an abrupt end. Alternatively, this conflictual interaction may also gradually de-escalate because the participants abandon their mutual attacks and return to a more rational and less offensive expression of different viewpoints. By examining conflictual interactions in a pharmacy patient consultation, Nguyen (2011) also noticed that even a long time after its resolution, the same conflict may be reinitiated with the previously accepted view automatically becoming the opposing view and the target of criticism. In addition, changes in intensity are not neat or homogenous. During the escalation of conflictual interactions, the conflictual linguistic strategies used by the current speaker may be less offensive than those used by the previous speaker. During the de-escalation stage, the conflictual linguistic strategies used by the current speaker may be more offensive than those used by the previous speaker. In other words, although from a holistic perspective, conflictual interactions show a consistent trend of escalation or de-escalation over a period of time, the intensity of words expressed by two adjacent speakers may fluctuate and be contrary to the overall trend during that period. This mismatch may be caused by the speaker's use of linguistic politeness to modify the intensity of conflict based on their interactional goals and personal styling. As studies on linguistic strategies suggest, the use of linguistic (im)politeness is likely to lead to different degrees of damage caused by the speaker's conflictual linguistic strategies to the recipient's face. However, studies on the overall

structure do not further investigate the relationship between this change in the FTAs and the fluctuation of conflict intensity or the escalation and de-escalation of the overall conflictual interaction.

2.3 Summary and conclusion

This chapter surveyed linguistic research on conflictual interactions. It shows that conflict studies are largely based on face-to-face conflictual interactions and analyse these interactions by using DA and (im)politeness theories. They investigate two broad issues. First, they identify conflictual linguistic strategies, explore the specific ways in which speakers express and investigate the impact of contextual factors (e.g., cultural, social, historical and demographic factors) on speakers' use of strategies and on how different conflictual linguistic strategies influence interlocutors' relationships. Second, they view conflictual interactions as a dynamic process and pay particular attention to the role of speakers' linguistic strategies in initiating a new conflictual interaction and in controlling an ongoing one. Follow-up surveys are commonly used in these studies to confirm the results of textual analysis and to obtain additional information that is unavailable from texts.

However, since most studies deal with conflictual interactions in face-to-face settings, it is still unknown whether the issues and topics they raise are applicable in the same way to conflictual interactions on Twitter and other CMC contexts. They typically focus on how speakers object to their opponents' points of view while few studies explore how speakers attack their opponents. The commonly applied three-stage view of conflictual interactions (i.e., onset, middle and end) places a great deal of emphasis on how speakers initiate and end such interactions. There is little attention given to the relationship between the nature of linguistic strategies and the structure of conflictual interactions. That is, little is known about the role of speakers' linguistic strategies in driving the development of conflictual interactions from beginning to end without interruption. These studies also ignore how the intensity of conflict changes as the interaction evolves, such as how conflictual interactions escalate and de-escalate and how the escalation and de-escalation fluctuate according to the use of linguistic (im)politeness. In order to reveal the developmental mechanisms of conflictual Twitter interactions that are constrained and influenced by Twitter algorithms, the present study, therefore, needs to address the following issues that are intimately related to this purpose:

(1) What linguistic strategies are used to (de-)escalate conflictual Twitter interactions?

(2) How do these strategies drive the development of conflictual Twitter interactions?

(3) How does linguistic (im)politeness function in conflictual Twitter interactions?

Since Twitter contexts are generally considered to be discourteous in contrast to face-to-face settings, this study is more interested in the issue of linguistic politeness. It intends to investigate whether Twitter users employ linguistic politeness when doing conflict, how they use it and why they do so. The third question is thus reformulated as follows: What is the most typical makeup of linguistic politeness on Twitter? How is linguistic politeness used to modify the intensity of conflictual Twitter interactions?

The review also found that methodological approaches in existing conflict studies appear to be appropriate for collecting and analysing online conflictual interactions. Two types of data are commonly collected: (1) the texts of conflictual interactions and (2) the follow-up surveys of individuals in relation to these collected interactions. Textual data faithfully records the entire process of conflictual interactions, including how they emerge, develop and terminate and provides detailed information about participants' language use and their interactive contexts. Corpus linguistics is useful for systematically approaching such data as it excels at sampling, collecting and pre-processing large amounts of naturally occurring linguistic data (McEnery, et al., 2006; McEnery & Hardie, 2013; Kennedy, 2014). Follow-up surveys help to obtain additional information that is not readily available from texts, such as people's attitudes and evaluations of online conflict and reasons for participating in conflictual interactions in particular ways, serving to enrich and verify the textual analysis.

In terms of data analysis, (im)politeness theories and DA, seem to provide suitable analytical frameworks for the textual data. (Im)politeness theory delves into the nature and makeup of people's language use. It helps to explain why the current speaker uses certain linguistic features or patterns to respond to the previous speaker in more detail and reasonably predicts the possible damage caused by the speaker's words to the recipient's face. DA takes a global perspective, exploring people's language use in discourse. It is concerned with the interrelationship between people's language use and contextual settings and also investigates how sentences and/or utterances are combined to produce texts and/or interactions. Combining these two frameworks allows an examination of the limitations and effects of Twitter algorithms on users' language use and interactional styles.

More importantly, it helps to identify speakers' conflictual linguistic strategies, including their makeup and nature and to explore, from a dynamic perspective, how these strategies drive the unfolding of conflictual interactions and change the intensity of such interactions. The next chapter further discusses these methodological approaches with the aim of establishing an analytical framework for the present study.

Chapter 3 Researching Suitable Approaches

3.0 Introduction

Chapter 2 reviewed linguistic research on conflictual interactions, including studies on conflictual linguistic strategies (i.e., disagreements and personal attacks) and studies on the interactional structure (i.e., onset, end and middle). It argued that in order to fully comprehend how conflictual interactions in CMC are formed and developed, the nature of linguistic strategies and the structure of such interactions need to be investigated together as the two are closely related. In the meantime, the impact of the algorithms of different CMC modes on the use of strategies and interactional patterns should also be considered. The discussion revealed three major tasks that need to be addressed, including the identification of linguistic strategies and linguistic politeness, as well as the analysis of their discursive functions. Chapter 2 also surveyed the methodologies used for collecting and analysing conflictual interactions in current research. It found that to address these three tasks, this study has to build a corpus of conflictual interactions and to conduct an in-depth textual analysis of the data in the corpus. The review suggests that a combination of DA and (im)politeness theories is suitable for systematically addressing all three issues. It also argued that follow-up surveys are a useful method for obtaining Twitter users' assessments and attitudes of the key findings of the textual analysis and for providing additional information that is not readily available from interactions alone.

This chapter discusses in more detail the premises of these analytical frameworks and approaches. The purpose is to adapt the analytical framework which is generally used in pragmatic studies based on face-to-face data to Twitter contexts and to integrate approaches from multiple disciplines including sociolinguistics and corpus linguistics into this classic framework. Specifically, this chapter critically reviews the current theories of corpus linguistics (see Section 3.1), sociolinguistic surveys (see Section 3.2), DA (see Section 3.3) and (im)politeness theories (see Section 3.4). It investigates the pros and cons of using these frameworks and approaches in the study and discusses the problems encountered in building a corpus of conflictual Twitter interactions and in adapting the analytical approaches designed for analysing face-to-face interactions to virtual interactions. It also puts forward possible solutions such as recommending that follow-up surveys may be integrated into the framework.

3.1 Corpus linguistics

The modern field of corpus linguistics emerged during the late 1950s (McEnery & Hardie, 2013). It provides reliable methodological support for collecting and processing a tremendous amount of naturally occurring linguistic data and has triggered significant changes in research in linguistics-related fields (e.g., McEnery et al., 2006; McEnery & Hardie, 2013). It has enabled linguists to collect massive linguistic datasets for at least some languages with a long writing tradition that was never thought possible before modern corpora appeared (Kennedy, 2014). The use of software for automatic annotation and analysis greatly reduces the time spent on investigating language data, improves research efficiency and minimises the need for manpower and material resources (Kennedy, 2014). Through the use of statistics, linguists are able to explore language phenomena empirically and visually by tracking the distribution of language items and to discover new phenomena that are difficult to detect in smaller datasets (Kennedy, 2014). Modern corpora also weaken the barriers between traditional disciplines and promote interdisciplinary communication (Kennedy, 2014). They provide traditional linguistic research with practical significance, bringing linguistic research closer to social life and accelerating the application of linguistic achievements in real life (McEnery & Hardie, 2013). This section discusses in depth the possibility of collecting Twitter data in a systematic manner based on corpus linguistics theories. It investigates whether current corpora are suitable for exploring the linguistic features in conflictual Twitter interactions and their interactional functions (see Section 3.1.1) and examines the challenges and possible solutions for building a corpus of such interactions (see Section 3.1.2).

3.1.1 Well-developed Twitter corpora

To study the overall structure of conflictual YouTube interactions, Bou-Franch and Blitvich (2014) built a sub-corpus with 23k words by selecting 444 comments on a YouTube video from the GenText Corpus, which is a specialised corpus designed to study gender inequality in Spanish and British societies with 42 million words (see Section 2.2 in Chapter 2). Specialised corpora are representative of a given type of text (e.g., fiction, newspaper, academic publications, lectures, online fora) and are limited to specific topics, subject areas and domains (e.g., medicine, economics, law, nature, environment), functioning to investigate a particular type of language use (Dehcheshmeh, 2007). Although there are many

specialised Twitter corpora, to my knowledge, none of them can be used to examine the developmental mechanisms of conflictual interactions in detail (e.g., Scheffler, 2014; Burnap & Williams, 2015; Imran et al., 2016; Waseem & Hovy, 2016; Davidson et al., 2017; Alorainy et al., 2018; Watanabe et al., 2018; Pflugmacher et al., 2020; Lundberg & Laitinen, 2020; Alharbi & Lee, 2021; Grimminger & Klinger, 2021; Masud et al., 2021). There are two possible reasons for this.

First, the linguistic features identified based on existing Twitter corpora are not generalisable because they tend to place explicit restrictions on the topics or types of data. For example, Imran et al. (2016, p. 1638) built a corpus of crisis-related posts on Twitter because Twitter has been widely used as "an active communication channel during mass convergence events, such as natural disasters" and has played an important role in decision making in "humanitarian and formal crisis response organizations, such as government agencies, public health care NGOs and military". They collected more than 52 million crisis-related posts on Twitter from 19 different crises that occurred between 2013 and 2015 in various geographical areas, thus ensuring that their dataset is sufficiently representative. However, the linguistic features identified in Imran et al.'s (2016) corpus may not be applicable to conflictual Twitter interactions around topics other than natural disasters. They are not suitable for the present study which intends to explore the linguistic features and dynamic nature of conflictual interactions on Twitter at a relatively general level. The dataset has to capture a representative range of conflictual Twitter interactions with a range of features and structural makeup. A case in point is Shum and Lee's (2013) dataset, although they did not construct a corpus. In order to study how users generally express disagreements on Hong Kong Internet discussion fora, they collected conflictual interactions derived from posts about highly and less controversial topics and did not further restrict the participants demographically or socially (see Section 2.1.1 in Chapter 2). This suggests that in order to guarantee that the dataset represents a representative range of conflictual interactions on Twitter, this study has to avoid limiting the participants, topics and types of interactions. Current corpora are, however, unsuitable for satisfying such data requirements.

Second, the interactional function of linguistic features cannot be investigated based on the majority of the current Twitter corpora as they generally collect individual posts rather than interactive data using the Application Programming Interface (API) provided by Twitter. API refers to "a set of programming code that enables data transmission between

one software product and another" ("Altexsof—What is API: Definition, types, specifications, documentation", 2021).[18] The Twitter API functions to quickly collect large volumes of posts, including tweets, retweets and replies, as well as metadata such as the time of posting and the number of times a tweet has been retweeted by other users (Scheffler, 2017). It can gather historical posts and posts immediately after being issued, by setting search parameters such as username, text content of tweet and the date of post, without soliciting user consent (Scheffler, 2017). However, it cannot retrieve and collect conversations sparked by the tweet (Zubiaga et al., 2016; Scheffler, 2017). This functional limitation may be responsible for the fact that there are few corpora designed for Twitter interactions. For the present study, in order to identify linguistic strategies and their discursive functions in conflictual Twitter interactions, the corpus should focus on conversations rather than isolated posts. However, most of the Twitter data currently available are individual posts in isolation and are therefore not suitable for this study.

Third, the corpora established by Zubiaga et al. (2016), Scheffler (2017) and Aktaş and Kohnert (2020) are the corpora of Twitter interactions that I could find so far but none of them are suitable for addressing the research questions of this study. The reply function is the key to interpersonal interactions on Twitter (Bruns & Moe, 2014; Zubiaga et al., 2016; Gruber, 2017; Housley et al., 2017; Scheffler, 2017; Aktaş, & Kohnert, 2020). Specifically, when Twitter users open the Twitter website or app, they are immediately confronted with a large number of posts. These posts may include original ones sent by Twitter users by clicking the 'tweet' button, known as *tweets* or *original tweets*, or quotations of other users' posts on Twitter by clicking the 'retweet' button, called *retweets* ("Twitter—How to tweet", 2022; "Twitter—How to retweet", 2022). Twitter users can access any post displayed on their screens and comment on this post by using Twitter's reply function to express their support or objections and to interact with the user who released this post. However, the conversations in Aktaş and Kohnert's corpus (2020) are incomplete. According to Scheffler (2017), the structure of Twitter interactions is tree-like. Each post may receive numerous replies, creating a multitude of interpersonal interactions that resemble one-to-many interactions (Scheffler, 2017). These replies may then be further commented on by multiple Twitter users, generating a complex network of the diverse types of online interpersonal

[18] Accessed, 22nd February 2022: https://www.altexsoft.com/blog/engineering/what-is-api-definition-types-specifications-documentation/

interactions (Scheffler, 2017). However, Aktaş and Kohnert (2020) only collected one-on-one interactions that contain the greatest number of replies of all interactions triggered by the same tweet. Zubiaga et al.'s (2016) and Scheffler's (2017) corpora do a better job addressing this issue. Zubiaga et al.'s (2016) corpus, however, includes only rumour-themed Twitter interactions, which lacks data about conflictual interactions arising from non-rumour or daily topics. Scheffler's (2017) corpus consists only of German language posts, which may be problematic for an in-depth analysis based on Brown and Levinson's (1987/1978) politeness theory. This chapter will return to this issue later in Section 3.1.2. Since current Twitter corpora are not appropriate for addressing the research questions, this study has to establish its own corpus. The next section discusses the processes, challenges and possible solutions in detail.

3.1.2 Building a Twitter corpus

According to McEnery et al. (2006), in order to establish a corpus, linguists must ensure that the data they are about to collect is efficiently representative to address their research purpose. They need to consider typically what constitutes a balanced range of data genres and a suitable sample size. Having collected data, linguists also need to annotate it to assist computers in reading and processing naturally occurring language data (McEnery et al., 2006). These two steps, however, are manifested quite differently, when building a corpus intended to serve different purposes.

The Twitter corpora established based on the above-mentioned guidelines generally serve the following two purposes: (1) to provide some standards for corpus construction and annotation or to provide comprehensive corpora for later analyses (Kennedy, 2014) and (2) to develop and/or optimise natural language processing (NLP) techniques such as computational classifiers and language detection models (Kennedy, 2014). It is common for these corpora to include a large volume of data, sometimes up to a few million posts on Twitter (e.g., Scheffler, 2014; Burnap & Williams, 2015; Imran et al., 2016; Waseem & Hovy, 2016; Zubiaga et al., 2016; Davidson et al., 2017; Alorainy et al., 2018; Scheffler, 2017; Watanabe et al., 2018; Pflugmacher et al., 2020; Lundberg & Laitinen, 2020; Aktaş & Kohnert, 2020; Alharbi & Lee, 2021; Grimminger & Klinger, 2021; Masud et al., 2021). Although researchers pay careful attention to the representativeness and balance of the data, these goals are usually satisfied through a large volume of data. In particular, the

improvement in the computing and storage capabilities of computers has directly contributed to the prevalence of the idea that representativeness can be easily achieved by increasing the amount of data (McEnery et al., 2006). To help computers to extract linguistic information, these corpora are usually annotated with linguistic features such as sentiment scores, POS information, the lemmatisation and word classes of lemmas and textual features such as paragraph or sentence boundaries, speech turns, line numbers and word breaks (McEnery et al., 2006; Kennedy, 2014). They are sometimes also encoded with contextual information such as geographic area and time, and extra-textual information like text types for written data (e.g., academic writing, online communication, letters) and the sociolinguistic variables of speakers for spoken data (e.g., gender, age, religion, social status) (McEnery et al., 2006; Kennedy, 2014).

To build a corpus that is used for DA, however, linguists often have very different considerations regarding the representativeness and annotation of data. Corpus based DA generally relies on two types of data. The first type involves a series of computational data collection and annotation methods and allows researchers to describe and explain the structure and distribution of naturally occurring language data and some specific issues quantitatively (e.g., language acquisition, variation and change) (e.g., Nguyen et al., 2013; Kim et al., 2014; Oz et al., 2018; Beach, 2019; Ross & Caldwell, 2019; Çöltekin, 2019; Reyes-Menendez et al., 2020; Harvey, 2020). The second type relies heavily on manual methods for collecting and annotating language data (e.g., Shum & Lee, 2013; Zhang & Kramarea, 2014; Iveson, 2017; Altoaimy, 2018; Bouko & Garcia, 2019; Lee, 2020; Hamdi, 2021). These corpora are designed to assist linguists in conducting an in-depth qualitative analysis of "the primary concern of the discourse or linguistic or social phenomena" (Ishihara & Prado, 2021, p. 643; see also Weisser, 2020). For example, identifying linguistic features or patterns that serve as a specific function and exploring the language use of a particular social group or in a specific social domain or context. Once these phenomena have been identified using qualitative methods, a quantitative analysis may be carried out based on the corpus.

Compared to the corpus used for developing and/or optimising corpus linguistic theories or NLP techniques, the size of the corpus used for DA is generally much smaller with less than about 10,000 words. DA requires researchers to delve into specific language use in contexts such as exploring in detail how social and personal factors affect language use and how messages are communicated and produce implications in interactions (Johnstone,

2018/2002). For a corpus designed for DA, data diversity or the range of data types is thus far more meaningful than the quantity of the data. However, the size of datasets for a quantitative analysis is generally larger than that for a qualitative analysis, due to the different data requirements.

In quantitative research, corpora are typically used to summarise and analyse the characteristics and distributions of pre-determined linguistic elements, whose annotation and analysis can be fully supported by computational technologies (e.g., Nguyen et al., 2013; Kim et al., 2014; Oz et al., 2018; Çöltekin, 2019; Beach, 2019; Ross & Caldwell, 2019; Harvey, 2020; Reyes-Menendez et al., 2020). These technologies have reached a reasonable level in automatically detecting and annotating various types of grammatical, lexical and syntactic information. For instance, in English-based datasets, the accuracy of the POS tagger is over 96% and that of the parser with crossing bracket measures is more than 92% (Hovy & Lavid, 2010). There are also a number of well-developed software packages and specialised websites designed specifically for corpus analysis, such as AntConc and Sketch Engine. They provide a variety of in-built tools that are suitable for a quantitative analysis of the corpus such as counting tokens, making concordances and producing n-grams, and also greatly speed up the process of data annotation and analysis. This contrasts with the case of corpora designed for a qualitative analysis.

In qualitative research, corpora serve to identify new features or patterns of individuals' language use, rather than quantifying the frequency of certain language phenomena or suggesting the most representative or distinctive linguistic feature(s) based on the frequency (e.g., Shum & Lee, 2013; Zhang & Kramarae, 2014; Iveson, 2017; Altoaimy, 2018; Bouko & Garcia, 2019; Lee, 2020; Hamdi, 2020). Ready-made annotation tools tend to treat variants in language use in Twitter interactions as noise to be removed such as misspellings, abbreviations, alphabetic character lengthening, superordination and unusual white space (Ilbury, 2020). These features, however, may play a very important role in a qualitative analysis, especially in identifying linguistic features in conflictual interactions. In addition, these tools are unable to handle more complex linguistic issues such as interpreting the pragmatic meaning of language use, categorising different types of discourses and analysing the meaning and functions of multimodal features (McEnery et al., 2006). As a result, corpora used for qualitative research are rarely annotated in advance.

Since this study aims to investigate a semantic entity, that is conflictual linguistic strategies, and their representations, pragmatic meanings and discursive implications, its data has to first be analysed qualitatively before being quantified. Quantification may include, for example, quantifying the frequency of different strategies in the corpus, calculating the proportion of replies including specific strategies out of the total number of replies and identifying at which stage certain strategies are more likely to appear. The dataset in this study should be appropriate for identifying conflictual linguistic strategies manually, without having to crunch a large amount of data. However, it cannot be too small in that it has to also include sufficient types of Twitter interactions that are representative of Twitter interactions to examine in a holistic manner how conflict arises and develops in most cases on Twitter. Thus, a central challenge of this study is to formulate a detailed Twitter data classification standard to guarantee that the corpus covers a sufficiently diverse and large set of interactions while also being feasible for a manual qualitative analysis. In terms of data annotation, although ready-made annotation tools are limited in this study, the annotation approach is still suitable for carrying out a comprehensive analysis of conflictual linguistic strategies.

A manual annotation approach is thus taken into account, which is typically used in research that aims to optimise automatic hate speech detection techniques, (e.g., Burnap & Williams, 2015; Imran et al., 2016; Waseem & Hovy, 2016; Davidson et al., 2017; Watanabe et al., 2018; Grimminger & Klinger, 2021). For example, Watanabe et al. (2018) built a corpus of hate speech on Twitter by merging three types of published datasets and created a labelling criterion for distinguishing 'hateful', 'offensive' and 'clean' content. They manually annotated some data based on this criterion and treated the manual annotations as a template. By iteratively comparing manual and computational annotations, they then trained and optimised the classification model. The key to success here is that this annotation/classification criterion is generally rational but Watanabe et al. (2018) did not explain in detail how they developed it. This issue is quite common in this type of research. Studies only briefly mention that they created a standard by adapting one or more currently available ones (e.g., Imran et al., 2016; Waseem & Hovy, 2016; Grimminger & Klinger, 2021). Sociolinguistic surveys may provide an effective method for ensuring that the self-created annotation and/or identification standard is reasonable and generally accepted. The next section critically reviews this method.

3.2 Sociolinguistic surveys

As a popular approach in applied linguistics, *sociolinguistic surveys* are widely used to investigate in depth individuals' linguistic behaviour in specified social environments and to explore the underlying reasons for these actions (Cooper, 1979). This method requires researchers to collect information about "the social organization of language behaviour and behaviour toward language in specified populations" (Cooper, 1979, p. 114). The purpose is to provide insights into the particular features of certain language use among individuals belonging to different social groups and in a variety of contextual settings (Cooper, 1979). Taking this approach, researchers are able to further test hypotheses regarding the relationship between social structure and language use or to address more practice-oriented concerns such as developing and evaluating language policies and promoting languages (Cooper, 1979). There are two primary techniques for collecting data in survey research: (1) sociolinguistic interviews and (2) questionnaires.

Sociolinguistic interviews refer to a method for gathering authentic "performance data" on language use and social interaction within specific contextual settings or social groups (Codó, 2008, p. 171). By restricting its informants and purposes, this approach provides researchers with a more controlled conversational environment than recording naturally occurring interactions (Codó, 2008). In this environment, researchers (i.e., interviewers) are able to develop an interpersonal interaction with informants (i.e., interviewees) based on specified purposes by constantly raising questions to the informants (Codó, 2008; Tagliamonte, 2012). Sociolinguistic interviews, however, are hardly used in conflict studies because linguistic data in such studies is generally obtained by observing interactions and the interview commonly serves quite different purposes. For example, Ben-Menachem and Livnat (2018) interviewed participants who participated in the interactions that they observed. This interview was used to verify Ben-Menachem and Livnat's (2018) identification of disagreement strategies in textual analysis and to obtain additional information that is unavailable in texts, including why speakers employ particular disagreement strategies and how they perceive the language uses of their opponents. This approach, however, is not suitable for analysing conflictual interactions on Twitter because the volume of participants in such interactions is relatively large and most participants are anonymous. Even if researchers interview Twitter users randomly, it is highly possible for them to fail to gather details of Twitter users' values, attitudes, beliefs and motivations

regarding conflictual interactions. This is because interviewees may be "reluctant to admit to negatively evaluated behaviour…in the presence of an interviewer" (Codó, 2008, p. 172). In this sense, questionnaires provide a more suitable resolution.

Questionnaires are another typical technique used in sociolinguistic surveys, which is more efficient for gathering information about sensitive issues (Codó, 2008). It is used extensively to obtain an overview of language use within a given social group or contextual setting (Codó, 2008). It is suitable for identifying target subjects from a large population sample and for gathering further quantifiable data about these targeted subjects' language uses and attitudes (Codó, 2008). Questionnaire surveys are also appropriate for providing additional information to complement other data collection methods because they allow linguists to collect discrete pieces of information without destroying or interfering with the naturally occurring language data (Codó, 2008; Boberg, 2017/2013). A case in point is Shum and Lee's study (2013). To verify their identification of disagreement strategies on two Hong Kong Internet discussion fora in a textual analysis, they interviewed users of these fora randomly. They also collected first-hand information about these users' interpretations of others' disagreement strategies in this survey. However, they did not explain why they did not invite the users who were involved in the interactions that they observed to participate in the survey. The present study argues that there are three possible reasons for selecting respondents in that way. First, participants on the Hong Kong Internet discussion fora are generally anonymous and the amount of them tends to be quite considerable. Although questionnaires may be distributed to the users of the observed interactions via the private message function provided by the platform, their response rates may be quite low. Second, surveying users in a broader sense helps to improve the universal nature of Shum and Lee's (2013) findings to some extent. It ensures that the users they observed and the users they randomly selected hold similar views about the use and expression of disagreement strategies. Third, this approach may preserve users' privacy. Shum and Lee (2013) did not reveal the names of these two fora but only mentioned that the selected fora were ranked 7th and 9th in the Hong Kong website according to the statistics provided by the Alexia company during the research period. Since Alexia's rankings are updated monthly, it is also unlikely to identify these two fora in the research.

The above discussions show that questionnaire surveys may help to address the issue that self-created annotation and/or identification standards lack credibility (see Section

3.1.2) because they are suitable for ensuring a greater level of the representativeness of a self-created annotation criterion. They also serve to provide additional information that is not readily available from the corpus alone, such as Twitter users' perspectives and attitudes. This point will be discussed further in Section 3.4. In Chapter 2, DA and (im)politeness theories were identified as the most suitable approaches for an in-depth qualitative analysis of the corpus to investigate the developmental mechanisms of conflictual Twitter interactions from a pragmatic perspective. The remainder of this chapter presents each of them in more detail.

3.3 Discourse analysis

The review in Chapter 2 shows that DA is suitable for investigating both the relationship between language use and contextual settings, as well as the relationship between language use and the overall structure of conflictual interactions. The concept of *discourse analysis* was first introduced by Zellig Harris (1952), who argued that discourse serves to provide insights into relationships across sentence boundaries and correlations between language and other behaviour. DA refers to a set of methods for analysing language use in discourse from two different perspectives (Paltridge, 2012/2006; Johnstone, 2018/2002; Jones, 2019/2012). It is used to examine the relationship between language and social and cultural contexts, in which language serves to integrate individuals into the social world through their representations of different perspectives and understandings and through their expressions of belongings to different social groups. Meanwhile, DA is also interested in the ways in which sentences and utterances are combined to realise texts and conversations. For the present study it was decided to follow Paltridge's (2012/2006), Johnstone's (2018/2002) and Jones's (2019/2012) classification, treating DA as an overall blanket term that compromise CA, despite the fact that the second focus overlaps with the typical focus of conversation analysis. The reason for this is that, in addition to speech production, this study is also concerned with the impact of Internet technology on interactions, which clearly falls outside the scope of hardcore CA. This section continues the discussion of DA. It first provides a brief overview of DA from the aspect of talk-in-interaction (see Section 3.2.1), then further explores how to use DA to explore conflictual linguistic strategies from a dynamic perspective (see Section 3.2.2) and finally discusses the challenges of using DA to analyse Twitter data (see Section 3.3.3).

3.3.1 A brief overview

Researchers employ DA to explore language-in-interaction by investigating how different turns are organised sequentially and, in particular, how speakers take turns to speak (Sidnell, 2010; Clayman, 2013; Hayashi, 2013). *Turn* originally represents "a time during which a single participant speaks, within a typical orderly arrangement in which participants speak with minimal overlap and gap between them" (Loos et al., 2013).[19] In DA, it typically refers to a stretch of words uttered by one speaker from start to finish, which is composed by a succession of turn-constructional units (TCUs) (Sidnell, 2010). Transition-relevance places (TRPs), which are formed at the time and place where TCUs may be completed, provide a venue where participants can decide the next speaker (Sacks et al., 1974; Sidnell, 2010; Clayman, 2013). Participants who assume the role of the current speaker may assign others to take the role of the next speaker by using various methods such as asking questions and using eye contact, gestures or facial expressions. If no speaker has been selected by the current speaker, any other speakers may self-select. In cases where no one has been selected or has self-selected, the current speaker may continue the conversation. DA argues that through these possible changes in speakerships, participants take turns to talk, thus driving the interaction forward.

Adjacency pair (AP) refers to "the unit of sequence construction that organises a great many actions in conversation" (Stivers, 2013, p. 192). Since DA claims that turns commonly emerge in pairs, one AP generally consists of two turns from different speakers (Goodwin & Heritage, 1990; Sidnell, 2010). The 1st and the 2nd turns belong to the class of the 1st pair part (FPP) and the 2nd pair part (SPP), respectively. The FPP and the SPP are organised by conditional relevance, meaning that the FPP determines the SPP and the SPP depends on the FPP because the SPP is generated on the basis of the FPP and reflects on the recipient's understanding of the FPP (Goodwin & Heritage, 1990; Sidnell, 2010). For example, in Excerpt 4, User A starts a phone call by introducing their identity, "This is Mr. Smith" (Line 01), and offering help, "may I help you" (Line 01). In response, User B accordingly accepts User A's offer, "Yes" (Line 02), and discloses their identity as User A, "This is Mr. Brown" (Line 02), thus providing the SPP for User A's turn. In Excerpt 5, probably due to signal problems, User B does not catch User A's words. User B indirectly asks User A

[19] Accessed, 22nd February 2022: https://glossary.sil.org/term/turn

to repeat the words by arguing that “I can’t hear you” (Line 02). However, since User B’s reply is irrelevant to User A’s offer, Lines 01 and 02 do not constitute an AP.

Excerpt 4 (Sacks, “Rules of Conversation Sequence”, cited in Sidnell, 2010, p. 11)

Line	Speaker	Text
01	User A	This is Mr. Smith may I help you
02	User B	Yes, this is Mr. Brown.

Excerpt 5 (Sacks, “Rules of Conversation Sequence”, cited in Sidnell, 2010, p. 11)

Line	Speaker	Text
01	User A	This is Mr. Smith may I help you
02	User B	I can’t hear you

In this regard, DA is an appropriate approach for studying the structure of conflictual interactions beyond the sentence level. It can be used to examine how the current speaker proposes their turn based on their understanding of the previous turn. It can also be employed to investigate how the current speaker’s turn is then interpreted by the next speaker and elicits the next turn. By using the turn-taking and adjacency pairs, the following section further discusses the relationship between conflictual linguistic strategies and the overall structure of conflictual interactions.

3.3.2 Discussion 1: turn-taking and developmental mechanisms

Conflict studies found that when the viewpoint of the previous speaker differs from that of the current speaker, the current speaker employs uninhibited linguistic behaviour to defend themself (e.g., Ben-Menachem & Livnat, 2018; Shum & Lee, 2013; Bou-Franch & Blitvich, 2014; Decock & Spiessens, 2017ab; Fernández-Amaya, 2019; Almutairi, 2021). This defence may involve one or some of the following strategies: objecting to the recipient’s viewpoint, belittling the recipient’s personal characteristics or depriving the recipient’s right to speak. From a DA perspective, before selecting linguistic strategies, the current speaker needs to interpret the previous speaker’s language use. Based on this interpretation, they then organise their turn by using conflictual linguistic strategies to respond to the previous speaker. The current speaker’s and the previous speaker’s turns thus constitute an AP and play the roles of the FPP and the SPP, respectively. Taking this approach, conflictual linguistic strategies are primarily identified based on the two-turn interaction fragment, which accounts for only a small part of conflictual interactions.

Harvey Sacks argued that the overall structure is "a relatively external source of interactional coherence that 'reaches into' sequence organization, turn construction and opportunities for participation" (Robinson, 2013, p. 258). This suggests that to reveal how conflictual Twitter interactions develop, an in-depth analysis of turn-taking within the entire interaction is necessary. Conflictual interactions should be taken as a series of turns that take place in a chronological order and their unfolding should be investigated step by step. For example, how do conflictual Twitter interactions arise from an AP? Having started, how do they develop through several turns towards the termination? The analysis should also examine how conflictual interactions, triggered by different opinions, progress as the participants exchange viewpoints. How does the current speaker respond to the previous turn and what responses does the next speaker make to the current speaker? At what stages of an interaction does uninhibited behaviour commonly occur and how does this happen?

Schmitt and Reiter's (2019) study, as well as Bou-Franch and Blitvich's (2014) work (see Section 2.2 in Chapter 2), is one of the few studies that explore conflictual interactions by analysing turn-taking. It examined how members of a German start-up team deal with a conflict about the selection of a company name over a six-month period. It adopted a relational continuity perspective on the basis of acknowledging "the dynamically emergent character of conflict as a process" (Schmitt & Reiter, 2019, p. 287). Taking this view, it investigated how language use in the current interaction is influenced by previous interactions and how it affects further conversations. However, similar to Bou-Franch and Blitvich (2014), Schmitt and Reiter (2019) only emphasised the beginning and end of conflictual interactions. The progress of conflictual interactions has not been explored in detail yet. Studies do not address the issues such as how the current speaker decides suitable uninhibited linguistic behaviour based on their understanding of the previous speaker's turn, how they perform this behaviour by using linguistic strategies, how their turn is then interpreted by the recipient and how this interpretation then affects the recipient's selection of linguistic strategies. What roles the current speaker's turn plays in overall interactions, for example whether it initiates de-escalation, is not investigated either.

DA is therefore a suitable approach for exploring speakers' linguistic strategies by examining whether the current speaker's turn satisfies the previous turn's requirements in terms of structure and content. However, APs and turn-taking are not sufficient for explaining why the current speaker designs their turn by using diverse conflictual linguistic

strategies combined with different forms of linguistic politeness. This suggests that a theory or method for exploring linguistic details within each turn is needed. This chapter will return to this point later in Section 3.4. Before using DA to analyse the corpus for this study, another matter to discuss is how DA can be applied to Twitter interactions because it was originally designed for face-to-face interactions. Since speech production on Twitter differs greatly from that in face-to-face settings, the next section explores how to employ DA to examine the development of conflictual Twitter interactions.

3.3.3 Discussion 2: turn-taking in Twitter contexts

With the rise of research on CMC, researchers have been adapting the DA frameworks to the special contingencies of various CMC modes such as multiparty chat rooms (e.g., Garcia & Jacobe, 1998, 1999; Schönfeldt & Golato, 2003; Benwell & Stokoe, 2006; Anderson et al., 2010; Jenks & Firth, 2013), instant messaging (Berglund, 2009; Meredith, 2017), newsgroups (e.g., Reed, 2001), online fora (e.g., Gibson, 2009; Stommel & Koole, 2010), Periscope (e.g., Licoppe & Morel, 2018), Soundcloud (e.g., Reed, 2017), FaceTime (e.g., Rintel, 2013), Email (e.g., Gruber, 1998; Condon & Čech, 2001), Skype (e.g., Licoppe & Morel, 2012) and Internet Relay Chat (e.g., Rintel & Pittam, 1997; Rintel et al., 2001). According to Giles et al. (2015) and Meredith (2019), since the technological affordances of CMC modes differ considerably, it is not clear that DA can be equally applied to analyse different online interactions without specific adaptations. Affordance refers to particular possibilities and constraints that CMC modes provide for interactions (Hutchby, 2001b). As far as I know, other than Housley et al.'s research in 2017, few interactional studies have adapted DA theories to Twitter contexts, that is to approach Twitter interactions as the data of DA.

The example used by Housley et al. (2017) in Excerpt 6 begins with a retweet from @mark_carrigan (i.e., Tweet 01). @mark_carrigan adds comments in the retweet to express their intention to use social media to improve student experiences. @EthicsWildfire asks @mark_carrigan: "why use a platform like Twitter for this over other mechanisms like internet portals etc" in Tweet 02. @mark_carrigan answers the question from Tweet 02 in Tweet 03. Then in Tweet 04, @EthicsWildfire acknowledges @mark_carrigan and raises a further question about the nature of Twitter, "does Twitter fit the typical UG demographic?". Based on these 4 tweets, Housley et al. (2017) argued that each individual

tweet serves as a TCU. @mark_carrigan and @EthicsWildfire facilitate the progress of the conversation by posting replies in turn. Each tweet functions as an interactive action (e.g., making an announcement, asking a question, acknowledging a response) and waits for a corresponding interactive action from the SPP as a response (Housley et al., 2017).

Excerpt 6 (Housley et al., 2017, p. 632)

Tweet	User	Text
01	@mark_carrigan	Once my book is all done & the set of following proposals are finished, I want to have a serious go at doing this The DDS Forum @DigitalSocSci Using social media to improve the student experience: creating a departmental back channel for undergraduates buff. Ly/IHwOw9y
02	@EthicsWildfire	@mark_carrigan A genuine, interested question: why choose a platform like Twitter for this over other mechanism like interact portals etc?
03	@mark_carrigan	@EthicsWildfire so facebook could also work. Just don't think building a propriety systems and inciting them to log in regularly will
04	@EthicsWildfire	@mark_carrigan Agreed – an easily accessible and accessed platform works well. However does Twitter fit the typical UG demographic? …

In Tweet 05, the post from @mark_carrigan functions as the continuation of their Tweet 03 and does not answer the question posed by @EthicsWildfire in Tweet 04. Housley et al. (2017) speculated that the occurrence of @mark_carrigan's Tweet 05 may be due to the temporal asymmetry between interlocutors. That is, @EthicsWildfire released Tweet 04 too early, and thus @mark_carrigan did not have time to answer @EthicsWildfire's question in Tweet 05. However, according to the reply from @EthicsWildfire in Tweet 06, Housley et al. (2017) finally argued that the temporal asymmetry between interlocutors has a minimal effect on the turn-taking organisation and attributed the occurrence of @mark_carrigan's Tweet 05 to Twitter's character limit. Based on this example, Housley et al. (2017) concluded that in Twitter interactions, users combine the norms of interactive sequences (e.g., turn-taking, recipient design, etc.) with Twitter's reply function to produce TCUs that unfold sequentially.

Excerpt 7 (Housley et al., 2017, p. 632)

Tweet	User	Text
05	@mark_carrigan	@EthicsWildfire not because I think Twitter is only platform that could be used for this purpose
06	@EthicsWildfire	@mark_carrigan TY. It's an interesting question of which existing platforms provide a ready made and popular mechanism for UG to use.

The study of Housley et al.(2017) does provide a valuable idea of approaching Twitter interactions as data of a DA analysis but the example used in this study is much simpler than

most cases on Twitter. For example, the above case shows a one-on-one interaction. On Twitter, however, it is more common to see that a single post triggers multiple replies and creates a one-to-many interaction, where more than one speaker 'talks' at the same time (Scheffler, 2017). If these replies are subsequently replied to by multiple users, Twitter interactions tend to resemble a tree-structure (Scheffler, 2017). This contrasts with the key to a successful face-to-face interaction that promotes "one party talking at a time" (Sacks, 2004, p. 37). Although it is also very common in face-to-face contexts that more than one speaker talks simultaneously, this practice is eventually amended to one speaker at a time. Housley et al.'s study (2017), however, neglects this one-to-many nature of Twitter interactions, which makes their explanation of the term 'turn' on Twitter problematic. Housley et al. (2017) argue that Tweet 05 from @mark_carrigan is the continuation of their Tweet 03. They, however, do not consider, from a structural perspective, whether @mark_carrigan's Tweet 05 is a direct reply to @EthicsWildfire's Tweet 04 or to @mark_carrigan's Tweet 03. The operations of these two cases on Twitter are completely different. In the former case, @mark_carrigan's Tweet 05 is published by clicking the reply function of @EthicsWildfire's Tweet 04. Tweets 03, 04 and 05 are shown in the same conversational stream. However, in the latter case, @mark_carrigan's Tweet 05 is issued by clicking the reply function of their own Tweet 03. This leads to Tweets 03, 04 and 05 being displayed in different conversational streams on Twitter: one includes Tweets 03 and 04 and another one includes Tweets 03 and 05. @mark_carrigan's turn here is divided into two posts, which challenges Housley et al.'s (2017) definition of TCUs. Therefore, before using DA to analyse conflictual Twitter interactions, this study still needs to examine the structure of Twitter interactions in more depth.

Overall, DA is suitable for exploring the developmental mechanisms of conflictual interactions in CMC because it not only focuses on speech production and distinctive turn-taking patterns but also pays attention to the impact of technical factors on human communication. However, it has limited value for further examining specific language design within each turn because turns are its smallest unit under investigation. DA tends to overlook the shaping of speakers' language use to discourse and the nuances of their language use in different contexts. A theory for investigating the design of turns is required in that the use of language within a turn is an essential part of the formation of an interaction. This study introduces (im)politeness theories, which conceptualise politeness as

a means of avoiding conflict and aggression in interactions, into the analytical framework. The next section explains (im)politeness theories, in more detail and pays particular attention to Brown and Levinson's (1987/1978) facework, which is regarded as "the only work that offers a systematic model for predicting what people will say in a given situation" by O'Driscoll (2019, p. 169).

3.4 (Im)politeness theory

Linguistic (im)politeness is, by far, one of the most common approaches in linguistics to provide a plausible explanation for nuances in language use. This section first outlines the historical development of major theories in (im)politeness research and critically discusses their strengths and weaknesses (e.g., Brown and Levinson, 1987/1978; Culpeper, 1996; Bousfield, 2008; Locher & Watts, 2005) (see Sections 3.4.1 & 3.4.2). It then elaborates why Brown and Levinson's (1987/1978) facework is more suitable for investigating the design of conflictual linguistic strategies by speakers in much more depth and also critically points out its shortcomings for the present study (see Section 3.4.3).

3.4.1 Brown and Levinson's (1987/1978) face-saving theory

Studies into (im)politeness can be traced back to the early days of linguistic pragmatic research, particularly the work by Austin (1962), Searle (1969) and Grice (1975), on what is unsaid in discourse (Haugh & Culpeper, 2018). The *speech act theory*, "a conceptual framework for describing and understanding the various kinds of linguistic action" (Sbisà, 2009, p. 231), was established by John L. Austin and his student John R. Searle. In this theory, *speech acts* are treated as the basic unit of human communication, serving to indicate the meaning of an utterance and consist of three variants: (1) locutionary acts (i.e., speech acts that are literally conveyed from the speaker's utterance), (2) illocutionary acts (i.e., speech acts that imply the true intention of the speaker's utterance) and (3) perlocutionary acts (i.e., speech acts that reveal the actual effect of the speaker's utterance on the recipient).

Similarly, Herbert P. Grice also noticed that the meaning in communication is very likely to convey more than the literal (or conventional) meaning of what words indicate. In order to explain how the recipient is able to recognise and interpret the speaker's unsaid meaning, Grice (1975) assumed that communication is an inferential process and

interlocutors are rational individuals who are "primarily interested in the efficient conveying of messages" (Fraser, 1990, p. 222). He maintained that rational efficiency prevails in an ideal situation of communication, which means that interlocutors are expected to convey what is informative, true, relevant and non-confusing. In other words, they are required to state something that is "neither more nor less than is required" (i.e., the quantity maxim), that is "genuine and not spurious" (i.e., the quality maxim) and that is "appropriate to immediate needs at each stage of the transaction" (i.e., the relation maxim) in an explicit and clear manner (i.e., the manner maxim) (i.e., Cooperative Principles; Grice, 1975, p. 45). However, in reality, the speaker usually behaves in a manner that deviates from the CPs. Grice (1975, p. 44) claimed that it is through the occurrence of such deviations that speakers are able to communicate "conversational implicature" and that the interpretation of these deviations serve to assist the recipient to infer the speaker's communicative intents. In the 1970s, a number of researchers such as Brown and Levinson (1987/1878), Lakoff (1973) and Leech (1977, 2014/1983) argued that linguistic politeness may serve as a type of conversational implicature. They found that speakers' desire to communicate politeness seems to lead them to diverge from the CPs during the course of a conversation. These "Gricean norm-based approaches" (Jary, 1998, p. 1) provided a solid foundation for the earliest systematic investigation into linguistic politeness in pragmatic research and gave rise to the first wave of research (Haugh & Culpeper, 2018).

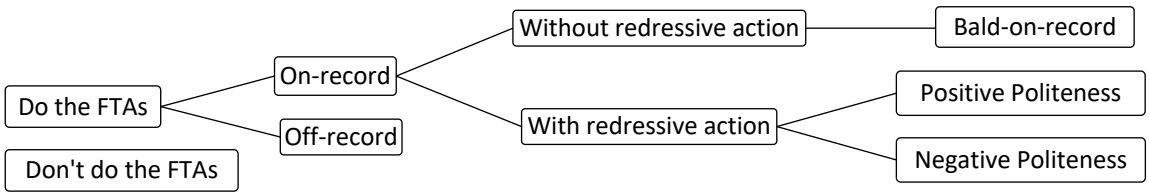
According to O'Driscoll (2019), Brown and Levinson's (1978/1987) face-saving theory is the most classic and also the most widely employed one in disagreement studies. They investigated linguistic politeness by formulating a framework of face, instead of introducing a principle for being polite and several maxims into Gricean theory, as Lakoff (1973) and Leech (1977, 2014/1983) did. Brown and Levinson (1978/1987) assumed that the rational actor of linguistic politeness is model persons (MPs), who speak natural language fluently and possess the special characteristic of face. They regarded *face* as individuals' intrinsic needs and defined it as "basic wants, which every member knows every other member desires and which in general is in the interests of every member to partially satisfy" (Brown & Levinson, 1987/1978, p. 62). Brown and Levinson (1987/1978, p. 62) then divided face into two sub-categories: (1) the *negative face* which is "the want of every competent adult member that [their] actions [are] unimpeded by others" and (2) the *positive face* which is "the want of every member that [their] wants [are] desirable to at least some others". They

assumed that “illocutionary acts inherently threaten either aspect of the face of another person” (Brown & Levinson, 1987/1978, p. 60), suggesting that every illocutionary act has the potential to harm the recipient’s face, either positive or negative face. This kind of illocutionary act was referred to as *face-threatening acts* (Brown & Levinson, 1987/1978). When the speaker ignores or negatively evaluates the recipient’s face-wants, desires, feelings, personal characteristics, beliefs, values or ideas, the recipient’s positive face is threatened; if the speaker’s speech interferes with the recipient’s freedom of action, the recipient’s negative face is at risk.

Brown and Levinson (1987/1978) subsequently proposed five “superstrategies” (Culpeper & Terkourafi, 2017, p.21) at various levels of indirectness and developed a flow chart to illustrate hierarchies of “output strategies” (Brown & Levinson, 1987/1978, p. 92) embedded in each superstrategy, as Figure 2 shows. This chart illustrates how speakers select suitable linguistic strategies to perform FTAs with varying degrees, including performing FTAs bald-on-record without redress, going on-record with redress (i.e., employing positive politeness strategies or negative politeness strategies), doing FTAs in an off-record manner and not doing FTAs. It can be found that linguistic politeness in Brown and Levinson’s (1987/1978) work is defined as a type of pragmatic strategy employed by the speaker to build/adjust/maintain the relationship between the recipient and themself by redressing their own FTAs. In addition, sociological variables were also integrated into Brown and Levinson’s (1978/1987) facework to further explain how speakers choose an appropriate strategy. They argued that in order to determine the extent to which the face-threat to the recipient may reduce, the speaker needs to assess the risks faced by the recipient in advance and to balance three sociological variables: (1) the social distance between the speaker and the recipient, (2) the relative power of the speaker with respect to the recipient and (3) the absolute ranking of impositions in the culture (e.g., the expenditure of goods and/or services by the recipient, the right of the speaker to perform the act and the degree to which the recipient welcomes the imposition) (adapted from Brown & Levinson, 1987/1978, pp. 71–83). Although both face-saving and face-damaging behaviour appears to be classified as facework, Brown and Levinson (1987/1978) examined only the previous one, which is designed to maintain the recipient’s face. In the mid-1990s, the first impoliteness theories emerged as a result of the lack of research on linguistic strategies that are used to reinforce threats to the recipient’s face (Culpeper & Hardaker, 2017; Haugh & Culpeper, 2018).

Figure 2

Possible strategies for doing FTAs from Brown & Levinson (1987/1978)



Note: adapted from *Politeness: Some universals in language usage*, by P. Brown and S. C. Levinson, 1987/1978, p. 69. Copyright 1987 by the Press Syndicate of the University of Cambridge.

3.4.2 Impoliteness theories

Impoliteness research was initially designed to study a mirror-image of Brown and Levinson's (1987/1978) work on linguistic politeness (e.g., Lachenicht, 1980; Culpeper, 1996; Bousfield, 2008). Researchers followed Gricean rational efficiency and acknowledged the concept of 'face' in Brown and Levinson's (1987/1978) face-saving theory (Culpeper, 2008, 2021). Being impolite in these early days, was generally treated as a type of intentional linguistic behaviour performed by the speaker in order to maximise the face damage to the recipient. For example, Jonathan Culpeper (1996, p. 350) defined linguistic impoliteness as "strategies that are designed to have the opposite effect [of linguistic politeness and that] are oriented towards attacking [the recipient's] face". He, subsequently, adapted Brown and Levinson's (1987/1978) framework to the case of impoliteness and identified five superstrategies, including (1) positive impoliteness (i.e., the use of strategies designed to damage the recipient's positive face wants), (2) negative impoliteness (i.e., the use of strategies designed to damage the recipient's negative face wants), (3) bald-on-record impoliteness (i.e., the FTA is performed in a direct, clear, unambiguous and concise way in circumstances where the face is not irrelevant or minimized), (4) sarcasm or mock politeness (i.e., the FTA is performed with the use of politeness strategies that are obviously insincere, and thus remain surface realizations) and (5) withhold politeness (i.e., the absence of politeness work where it would be expected) (adapted from Culpeper, 1996, pp. 356–358). Derek Bousfield is another influential researcher in the earliest impoliteness studies. In 2008, he published *Impoliteness in Interaction*, the first monograph in this area (Culpeper & Hardaker, 2017). Bousfield (2008, pp.72) regards impoliteness as a type of linguistic strategy

that functions to "constitute the communication of intentionally gratuitous and conflictive verbal face-threatening acts which are purposefully delivered" and developed a more comprehensive model of face-attacking strategies based on Culpeper's (1996) impoliteness framework. It is important to note that the early impoliteness studies generally adopted Brown and Levinson's concern about relational work, considering how the use of (im)politeness strategies relates to sociological variables. However, they still sought to identify 'authentic' impoliteness by capturing the speaker's intention (Culpeper, 2008, 2021; Locher & Watts, 2008). A growing body of research into intentionality indicates that (e.g., Gibbs, 1999; Haugh, 2008) intention does not seem to exist in the speaker's mind prior to their choice of (im)politeness strategies, contrary to what Gricean approaches commonly assume (Culpeper, 2021; Haugh & Culpeper, 2018). Researchers find that even when the speaker does not intend to damage the recipient's face, the recipient is still very likely to perceive the speaker as being impolite. Moreover, in accordance with relevance theory, a great number of studies (e.g., Watts, 1989, 2003, 2005/1992; Fraser, 1990; Jary, 1998; Eelen, 2001; Mills, 2003; Locher, 2004, 2006ab; Locher & Watts, 2005, 2008) also demonstrate that linguistic (im)politeness is not a relatively fixed behaviour for the recipient, since their interpretation of the speaker's intention appears to emerge through interactions (Haugh & Culpeper, 2018).

Relevance theory was proposed by Sperber and Wilson in 1986 as an alternative to Gricean norm-based approaches to communication. It claims that "what guides the interpretation of communicative behaviour is not the adherence to a set of conventions, but the way human central cognitive systems have evolved" (Jary, 1998, p. 3). Sperber and Wilson (1995/1986, p. 262) regarded communication as an ostensive-inferential process, during which "human cognition tends to be organised so as to maximise relevance". This means that interlocutors tend to select their stimuli in the light of relevance considerations and to communicate an assumption of optimal relevance by ostensibly utilising these stimuli (Jary, 1998). Specifically, in a conversation, the speaker communicates their intention through the discourse and allows the recipient to interpret it; as a consequence, the recipient generates a set of assumptions based on the literal meaning of the speaker's language use and the context in which it is used and then selects the most likely assumption based on a comparison between meaning and context. In response to his post-Gricean theory, several researchers (e.g., Watts, 1989, 2005/1992; Fraser, 1990) have explored a

relevance-theoretical account for linguistic (im)politeness, as well as the relationship between (im)politeness and markedness (i.e., linguistic behaviour is more likely to be noticed if it does not adhere to the social constrains of the ongoing interaction). They, however, did not delve into the pragmatic mechanisms by which the speaker's (im)politeness is recognised as 'unmarked' by the recipient in much detail (Cupeper & Terkourafi, 2017).

In 1998, Mark Jary detailed the inferential paths for the processing of polite expressions (Cupeper & Terkourafi, 2017). He followed Brown and Levinson's (1987/1978) assumption that the use of linguistic strategies is greatly influenced by people's assumptions about the relationship with each other. Jary (1998) also claimed that the interpretation of politeness is an exploration of the recipient's expected social relations with the speaker by observing the speaker's language use. The recipient has a prior assumption about their relationship with the speaker based on social culture and is subsequently forced to create a new one in response to the speaker's language use and contextual settings. From this relevance-theoretic perspective, linguistic politeness is only noticed by or communicated to the recipient when their new assumption is incompatible with their previous one. This cognitive bias indicates that the speaker holds the recipient in higher esteem than the recipient previously expected, which further leads the recipient to modifying their perception of the relationship. Jary (1998) argued that it is only in this case that an implicature of being polite is relevant enough for the recipient to process, which is, however, criticised by Marina Terkourafi. Terkourafi (2005) contends that Jary's (1998) work is not fully in accordance with a discursive approach because he still treated face as a static entity, pre-existing the interaction and unaffected by the interaction. Although these early relational studies did not provide a widely accepted framework for analysing (im)politeness, they indeed promoted the popularity of communicative approaches in this field in the 2000s, leading to the second wave of research (Haugh & Culpeper, 2018).

Relational work, also known as post-modern work, offers an alternative account for linguistic (im)politeness by employing discursive/communicative approaches and has significantly promoted the development of impoliteness studies in the last decade (e.g., Eelen, 2001; Mills, 2003; Watts, 2003, 2005/1992; Locher, 2004, 2006ab; Locher & Watts, 2005, 2008). It criticises Gricean norm-based approaches for accepting norms as the starting point and for establishing a "pseudo-scientific theory" (Culpeper, 2008, p. 19). Even though

the analysis of interlocutors' faces is still considered integral to relational work, researchers insist that they follow Goffman's (1967, p. 5) concept of face, which refers to "the positive social value a person effectively claims for themself". Face here is regarded as a matter of individuals' self and its relationship to surrounding contexts, rather than a reflection of people's psychological needs as stated by Brown and Levinson (1987/1978). Being (im)polite is thus considered to be a form of relational work. Researchers believe that (im)politeness cannot be defined and described in one simple and fixed way and its very concept is subject to discursive struggles in the course of ongoing interaction, reflecting the social struggle between discursive power and appropriateness. In contrast, the classic theories of (im)politeness focus only on how interlocutors' language use fits the concept of (im)politeness developed by theorists without taking into account communicators' own interpretation of being (im)polite and their struggles for the right to speak. In order to address this issue, relevance work employs discursive approaches and focus on the layperson's linguistic behaviour and perception, investigating how interlocutors negotiate their relationships with each other during interaction and also the fluidity of power in it. In addition, classic approaches developed a speaker-oriented framework with a strong emphasis on the speaker's language production and their subjective intentions embedded within their language use. Relational studies, however, have generally demonstrated that it is the recipient's interpretation of the speaker's intention rather than the speaker's intention per se that determines whether the speaker is polite or impolite. As part of this judgement, the recipient also needs to consider the relational aspects of the speaker's utterances in the light of the norms and expectations that they have constructed and acquired through their own experiences of similar situations. The belief that (im)politeness is unpredictable has become increasingly popular as a result (e.g., Watts, 2003; Graham, 2007, 2019).

Such relevance-theoretical accounts have been challenged in recent years, (Haugh & Culpeper, 2018). Firstly, although discursive approaches are suitable for inquiring into the interactivity between interlocutors and the fluidity of their discursive powers, they are nearly incapable of identifying and generalising specific linguistic patterns in interaction (e.g., Holmes & Schnurr, 2005; Terkourafi, 2005). For the purpose of the present study, relevance work may not be appropriate for investigating the role of linguistic (im)politeness in constructing conflictual linguistic strategies and its primary composition. Secondly, discursive approaches have also been challenged for overlooking the significance of

researchers' analysis and paying too much attention to lay understandings. Relevance work provides laypersons with the opportunity to account for their language use in a position of authority. Haugh and Culpeper (2018) assert that lay understandings may be recast in this approach as a formal theory with a function similar to that proposed by theorists. In relevance work, research focus is exclusively on interlocutors' discursive struggles around politeness and impoliteness. Communicative approaches, however, are relatively limited in their explanations of linguistic phenomenon referred to as (im)politeness, since they are insufficient to investigate in great depth certain situations in which participants and researchers jointly negotiate and co-construct interactions. The development of technology, for example, has been shown to influence and greatly alter human communication in virtual contexts. Researchers assume not only the role of being a developer of online communication technologies, but also that of an observer in interaction, functioning to interpret, intervene and reshape CMC users' linguistic behaviour. In addition, Locher and Watts (2008, p. 78) maintain that in relational work, the recipient's evaluation of the speaker's linguistic behaviour represents their own "cognitive conceptualization of forms of appropriate and inappropriate behaviour". It is evident from this discussion that communication is regarded as an ostensive-predictive process that evolves as a result of an increase in the amount of information received. However, the purpose of the present study is not to identify which of these two theories better captures the true nature of human communication, but rather to reveal the developmental mechanisms of conflictual Twitter interactions. It is concerned more with how to rationalise interlocutors' linguistic behaviour than with how they generate opinions in their minds nor how they perceive others' language design. This requires that researchers approach conflictual Twitter interactions from a logically rational perspective by assuming the role of an observer, not placing an excessive focus on individual differences. While this logic may sometimes be incongruent with interlocutors' own cognitive processes, it constitutes a plausible explanation from a third-party perspective. Therefore, although Brown and Levinson's (1987/1978) classic approach has become obsolete in (im)politeness studies and has often been criticised as unsuitable for analysing inharmonious language use, it still appears to be partly applicable for addressing the research questions in this study because it highlights researcher's analysis and pragmatic mechanisms in interaction. This results in a critical discussion of Brown and Levinson's (1987/1978) framework in the next section.

3.4.3 Discussion and criticism of facework

Brown and Levinson (1987/1978) adhered to Gricean norms, maintaining that interlocutors' linguistic behaviour is constrained by pre-existing conventions, such as social norms and tacit agreements. Interlocutors are believed to use a "rational means-ends reasoning" (Brown & Levinson, 1987/1978, p. 7) in their minds when interacting; this reasoning assists them in "mapping out logical, rational choices in order to achieve particular goals" (Culpeper & Terkourafi, 2017, p. 21). Discursive interpretation here refers to a process in which the recipient compares the literal meaning of the speaker's discourse with the conventions provided by the context and constantly seeks a reasonable explanation for the speaker's deviation from the conventions, thus further inferring the speaker's intention. However, as Section 3.4.2 suggests, this approach has been widely criticised for overemphasising the speaker's intention to use linguistic politeness (e.g., Eelen, 2001; Mills, 2003; Watts, 2003, 2005/1992; Locher, 2004, 2006ab; Locher & Watts, 2005, 2008). In 1987, Brown and Levinson clarified this issue, arguing that what they meant by the speaker's intention is actually the communicative intent that the speaker uses to convey a message to the recipient through language, not necessarily the speaker's innermost ideas. The recipient and other third parties are capable of reasonably inferring the speaker's communicative intent, based on how the speaker expresses themself when being polite (Brown & Levinson, 1987/1978). This clarification indicates that in an interaction, what passes between interlocutors is the message that they display externally and the inferences that speakers draw based on them, rather than a reflection of their psychological activities, although this message, in some cases, may be used to gain insight into interlocutors' psychological states. In other words, in Brown and Levinson's (1987/1978) perspective, linguistic politeness is conscious but does not necessarily equate to actual cognition. It is simply a linguistic strategy employed by the speaker to manage the possible threats posed by their utterances to the recipient's face. In terms of conflictual interactions, speakers may not always issue a turn in order to hurt the opponent's face or to force the opponent to accept an opposing viewpoint; they may simply wish to demonstrate a different point of view, but this action may, in reality, pose a threat to the opponent's face. This results in Brown and Levinson's (1987/1978) notion of communicative intent providing an appropriate approach for the present study, as it enables researchers to reasonably deduce speakers' communicative intent based on their language use in appearance.

In order to research the speaker's use of linguistic politeness, Brown and Levinson (1987/1978) developed the concept of facework. They used the word 'face', which originally refers to the surface of objects, to refer to rational desires of individuals about their relationship with others in an interaction. It functions to lead one party to draw inferences about deeper things of the other party (O'Driscoll, 2017). In other words, face refers to "something intrinsic to the nature of human beings when they interact with each other" (O'Driscoll, 2017, p. 91). Brown and Levinson (1987/1978), subsequently, argued that linguistic behaviour should also have a certain inherent property, that is the potential to damage the recipient's face. Facework, thus, involves a response to the potential properties of people's linguistic behaviour, as well as their own face needs. One important criticism of facework derives from its adaptability in different cultures. It is argued that facework is mainly applicable to Anglo-Saxon cultures, especially English-speaking cultures. The face-saving view of politeness is based on the assumption of MPs who are "unconstrained by social considerations and thus free to choose egocentric, asocial and assertive interaction" (Werkhofer, 1992, cited in Shahrokhi & Bidabadi, 2013, p. 23). However, some researchers found that in certain non-western cultures where group norms and values predominate over individuals' social behaviour, face protection seems to be less instrumental and plays a similar role to that of appropriate behaviour in maintaining social harmony (e.g., Lebra, 1976; Wierzbicka, 1985; Matsumoto, 1988; Ide, 1989; Gu, 1990; Nwoye, 1992; Pandaripade, 1992; Sifianou, 1993; Mao, 1994; Koutlaki, 2002; Eslami, 2005; Yang, 2010). Although these studies do indicate "cross-cultural differences in the standard (context-free) values attached to certain acts", O'Driscoll (2017, p. 98) asserts that they are "somewhat misguided". He argues that what Brown and Levinson's (1987/1978) facework depicts is actually the behaviour before it has been performed.

In face-saving theory, there is no statement arguing that the purpose of using politeness strategies is to communicate politeness. Brown and Levinson (1987/1978) added an additional layer of assumption to the statement, suggesting that every linguistic behaviour could potentially damage the recipient's face. It should be noted that Brown and Levinson (1987/1978) do not mean that linguistic behaviour necessarily results in face damage; they rather presupposed that linguistic behaviour has an intrinsic nature that may either be manifested or concealed. O'Driscoll (2017, p. 98) points out that FTAs are just "the 'worst possible reading' of an act". Brown and Levinson's (1987/1978) assumption is made

about linguistic behaviour in advance, not about "the intention of the speaker and/or the effect of the [linguistic behaviour] after it has been performed" (O'Driscoll, 2017, p. 98). They presented facework as a formula, a model for accounting for linguistic behaviour patterns of participants, which allows a wide range of possibilities for interpretation by speakers, recipients and also researchers. In this sense, facework is very flexible, as it is able to simultaneously address "reactive and proactive, supportive and antagonistic interpersonal phenomena, intention and effect, and behaviour addressed to self, other(s) or both" (O'Driscoll, 2017, p. 100). This means although facework is speaker-centred focusing on explaining what the speaker utters, its emphasis on the instrumental nature of rationality leads to an overwhelmingly consideration of the recipient's face (O'Driscoll, 2017). In the case of Twitter, the prevalence of uninhibited linguistic behaviour has led to the labelling of Twitter as a virtual space full of verbal violence. As a result of this label, Twitter users tend to arrive at a consensus that face-threatening acts are becoming the norm in Twitter interactions; this in turn may further influence and change people's perceptions and selection of linguistic strategies when interacting on Twitter. FTAs, which are derived from Brown and Levinson's (1987/1978) hypothesis, are very likely to become an established norm on Twitter. In this respect, Brown and Levinson's (1987/1978) face-saving theory appears more suitable for analysing conflict. It suggests that linguistic politeness may play an important role in adjusting the intensity of conflictual interactions, such as mitigating or even intensifying the conflict between interlocutors.

Although many criticisms of Brown and Levison (1987/1978) regarding cultural adaptability are due to a misperception of facework, it cannot be denied that culture plays a significant role in accounting for specific linguistic strategies through the use of facework. There is still a need for additional attention from researchers on cultural adaptability of facework when it comes to specific issues. For the present study, this point does pose an issue for the selection of Twitter data because the interpretation of linguistic politeness may vary between English and non-English conflictual Twitter interactions. This requires this study to focus on conflictual Twitter interactions in English although Twitter users interacting in English may not be native English speakers. While it may be possible to identify Twitter users' cultural background according to their expressions in English, this study argues that such identification is not necessary. On the one hand, there is the fact that Twitter users often hide their real-life identities or construct virtual identities which is likely to hinder the

identification of their cultural background (Ilbury, 2020). Researchers are unable to determine whether the linguistic features in English posts that do not match typical native English speakers' language use are necessarily reflective of their speakers' status as a non-native English speaker or whether they were intentionally designed by a native English speaker. So-called native English speakers also do not constitute a linguistically homogeneous group, as English is spoken across many different countries and its speakers belong to a range of social and regional groupings. On the other hand, being able to converse in English may indicate that the user has some knowledge of English etiquette. The cultural bias of Brown and Levinson's (1987/1978) facework does not then represent a major issue when studying Twitter interactions written in English.

Brown and Levinson (1987/1978) subsequently turned to linguistic strategies in their discussion, which includes an important drawback of their face-saving theory. They claimed that linguistic politeness is a type of linguistic strategy that serves to perform facework, especially the task of reducing threats to the recipient's face. They also rationalised various strategic choices of people by expounding comprehensively the impact of three sociological variables, including Distance, Power and Ranking of the imposition. There is no doubt that Brown and Levinson (1987/1978) encountered a logical issue when connecting their statements of facework with those of politeness strategies. They constructed facework as a flexible analytical framework that allows analyses of both the exacerbation and mitigation of face damage. During their discussions of linguistic politeness, however, the scope of facework was narrowed to face maintenance, which is why in Brown and Levinson's model, "politeness and facework are essentially the same thing" (O'Driscoll, 2017, p. 96). This leads to the fact that although facework is quite suitable for investigating the fluidity of language use in different contexts, Brown and Levinson's (1987/1978) analysis of politeness strategies reveals a crude simplification of this communication process. They examined only the speaker's language use, which generally takes the form of politeness and is employed when the speaker intends to protect the recipient's face. Following this, Brown and Levinson (1987/1978) posited that there is a one-on-one relationship between linguistic forms/strategies, their communicative intent and their communicative impact. However, does the use of polite forms necessarily indicate that the speaker's communicative intent is to mitigate? Is the use of impolite forms inherently indicative of the speaker's communicative intent to reinforce? Is it possible to use polite forms to reinforce face threats

and impolite forms to mitigate face damage? Brown and Levinson (1987/1978) only allowed one possible divergence from the one-on-one relationship: irony. They treated it as an off-record strategy that contains an explicit feature of politeness but also as a hint of the real intention of being impolite (e.g., clues may be prosodic, kinetic or simply contextual). In Leech's (2014/1983) politeness theory, irony functions to demean the recipient and assists the speaker to greatly avoid being accused of offending because the speaker can always defend their offence by claiming or implying that they are innocent.

Leech (2014/1983) attributed the 'understatement' and 'attitude clash' as the two triggers of irony. *Understatement* refers to an approach by which the speaker "makes a claim that is clearly inadequate, less informative than required" (Leech, 2014/1983, p. 237). Brown and Levinson (1987/1978) also mentioned understatement but only as an off-record strategy often used in conjunction with irony. *Attitude clash*, which was not addressed by Brown and Levinson (1987/1978), refers to cases "where the overt 'polite' meaning and the 'impolite' meaning of irony occur side by side in the same piece of language" (Leech, 2014/1983, p.238). To explain it, Leech (2014/1983) cited an example of what Culpeper (2011a, p.174) in his study of impoliteness referred to as "verbal formula mismatches": "Could you just fuck away?". The phrase of "Could you just" introduces a polite and highly mitigated request while the last two words "fuck away" expressing 'go away' are in the rudest manner. Leech (2014/1983, p.238) argued that by using an attitude clash, irony is able to be interpreted as polite before "being forced to retrospectively reinterpret…as ironical in the context of what follows". Graham (2019, p. 320) also noticed an attitude clash in conflictual interactions in Raihnbowkidz's video stream, "Please kill yourself you little slut." and explained it as an attack on Raihnbowkidz (see Section 2.1.2 in Chapter 2).

In addition, Leech (2014/1983) suggested an alternative possibility that the current speaker's overt impoliteness leads to a polite interpretation, which was not mentioned by Brown and Levinson (1987/1978) and most impoliteness studies. Leech (2014/1983, p. 241) called it *banter* and explained that "banter tends to occur almost ritualistically in certain communities of practice where individuals undergo emotional pressure", such as in hospitals or in sports team events, serving to promote an "atmosphere of friendly jocularity". Leech (2014/1983) cited an example from Mill (2003) that took place in a male-dominated teaching department:

> "CHRIST ALMIGHTY (.) if one man could bugger me around more in two days than you have in the last two days I'd like to know his name Christ Almighty (.) Still all right. If I didn't know you I would say it was on purpose." (Mills, 2003, p. 107)

Leech (2014/1983, p. 241) illustrated that although this turn could easily be interpreted as an explosive expression of anger without an obvious sign of banter, since it was accompanied by laughter and then received the recipient's response of "have a drink", it clearly functions as a joke. Brown and Levinson's (1987/1978) politeness study and most impoliteness studies, however, argued that language use taking a polite or impolite form suggests the speaker's communicative intent to preserve or hurt the recipient's face, respectively, and functions to mitigate or intensify the FTAs, respectively. They paid little attention to polite expressions that are used as aggression and impolite expressions that are used as friendliness. So how can this study adapt Brown and Levinson's (1987/1978) face-saving theory to explain the situations where the superficial meaning of speakers' turns is opposite to their communicative intent? Is it possible to address this issue by changing the framework of Brown and Levinson's (1987/1978) politeness strategies by abandoning the one-on-one relationship between politeness forms and their communicative impact? These questions need to be further explored through data analysis.

In summary, this review suggests that Brown and Levinson's (1987/1978) facework is appropriate for combining it with DA to conduct an in-depth pragmatic analysis of the makeup, nature, function of language use within each turn and the relationship between turns. It helps to examine a broader range of language use such as how the speaker realises their oppositions to the recipient and to analyse the extent to which subtle changes occur in conflictual interactions. This study argues that researchers are supposed to analyse the speaker's use of linguistic politeness from an interactive and dynamic perspective and to distinguish the speaker's and the recipient's understandings in the analysis. On the one hand, researchers should examine the speaker's communicative intent by referring to their use of linguistic features. On the other hand, researchers should also examine how the recipient expresses themself in their turn to determine how they perceive their (previous) speaker's communicative intent. This is because according to DA, in an AP, the way in which recipients design the SPP (i.e., their own turns) reflects their understanding of the FPP (i.e., the previous turn) (see in Section 3.2.1). However, neither Brown and Levinson's (1987/1978) facework nor DA are appropriate for explaining why the recipient does not

interpret the (previous) speaker's communicative intent in full accordance with the (previous) speaker's language use. It is likely to address this issue by conducting an in-depth pragmatic analysis to identify the different discursive functions of the speaker's and the recipient's understandings in the interaction. The questionnaire approach that has been widely employed in disagreement studies may provide another suitable solution (e.g., Shum & Lee, 2013; Ben-Menachem & Livnat, 2018; see Section 3.2). By surveying the recipient, this study can obtain information about how the recipient interprets the (previous) speaker's communicative intent in detail, as well as other information that is not readily accessible from a textual analysis, thus avoiding subjective and biased results.

Chapter 4 Methodology

4.0 Introduction

The previous chapters focused on identifying the nature of conflictual interactions in CMC and the methods used for studying it. Specifically, Chapter 2 surveyed linguistic studies on the structure of conflictual interactions focusing on disagreements and personal attacks. The review suggests that an investigation of the developmental mechanisms of conflictual Twitter interactions needs to delve into linguistic strategies and linguistic politeness and their discursive functions. The review also established that these aspects are generally examined using a combination of multidisciplinary frameworks and approaches including corpus linguistics, DA, Brown and Levinson's (1987/1978) facework and sociolinguistic surveys. Chapter 3 explored these approaches and frameworks including the kinds of adjustments needed for studying conflict in Twitter contexts. The review suggested that, apart from building a corpus of conflictual Twitter interactions and analysing the corpus by using DA and Brown and Levinson's (1987/1978) facework, it is also necessary to integrate short questionnaires to obtain first-hand information about Twitter users' understanding of Twitter interactions and their attitudes and/or evaluation of language use in online conflict.

In order to apply the identified frameworks and approaches to online interactions, three main adjustments should be made. First, corpus linguistics has been successfully used in CMC studies to develop and optimise computational language detection and/or analysis software and to quantify specific linguistic issues such as counting the frequency of linguistic features (e.g., Watanabe et al., 2018; Ross & Caldwell, 2019; Reyes-Menendez et al., 2020; Grimminger & Klinger, 2021; Masud et al., 2021). However, these tools are rarely used to build corpora for a qualitative analysis, especially the identification of linguistic strategies and the exploration of their interactional functions. This study should, therefore, establish a data classification and annotation schema to ensure that the dataset is suitable for a qualitative analysis. Second, since cases where the surface meaning of the speaker's turn is opposite to the speaker's communicative intent are quite common on Twitter, such as in the case of irony, Brown and Levinson's (1987/1978) facework does not, on first sight, appear to provide an adequate model of analysis. However, given that their subsequent analysis of linguistic politeness simplified the relationship between linguistic forms/strategies and communicative intents and impacts, it seems suitable to deal with these issues. This though

necessitates the identification of an appropriate way in which the facework can be better employed to investigate linguistic politeness in Twitter conflictual interactions. Third, although the DA framework has been adapted to interactions on various online platforms (e.g., Berglund, 2009; Anderson et al., 2010; Licoppe & Morel, 2012; Rintel, 2013; Meredith, 2017), to date there has been little systematic research on how to employ DA to investigate Twitter interactions. To facilitate the adaptation of the DA framework to Twitter contexts, pilot studies on the structure of Twitter studies are necessary.

This chapter further discusses how these approaches are employed in this study. Specifically, the chapter is structured as follows: Section 4.1 provides a brief overview of the questions that need to be addressed in this study and outlines the methodological approaches that are used to address these issues. Section 4.2 discusses the construction of the Conflictual Twitter Interaction Corpus. The section first examines how Twitter algorithms structure interactions and influence the nature of Twitter interactions based on several pilot studies (see Section 4.2.1) and creates a detailed Twitter data classification standard to ensure that a representative range of conflictual Twitter interactions is captured (see Section 4.2.2). It then presents in detail a set of procedures and methods for collecting conflictual Twitter interactions and the makeup of the corpus (see Section 4.2.3). It also establishes a proper annotation standard and verifies its suitability by designing and conducting a questionnaire suitable for people with and without linguistic training (see Section 4.2.4). Section 4.3 discusses the analytical methods in detail, designs a questionnaire for Twitter users about their attitudes and opinions regarding conflict on Twitter and outlines the ways in which DA, Brown and Levinson's (1987/1978) facework and the use of questionnaires can be integrated to analyse conflictual Twitter interactions. Finally, Section 4.4 presents some ethical considerations that govern data collection and data analysis on Twitter.

4.1 A brief overview of research questions and methodological approaches

In order to reveal the developmental mechanisms of conflictual Twitter interactions, this study needs to address two major questions (see Section 1.4 in Chapter 1): (1) how do people employ linguistic strategies to initiate and progress conflictual interactions on Twitter? (2) How do Twitter algorithms affect peoples' ideologies of language through constraining and shaping distinctive ways of speaking and communicating on Twitter? In order to answer these questions, the following issues need to be explored (see Section 2.3 in

Chapter 2): (1) what linguistic strategies are used to (de-)escalate conflictual Twitter interactions? (2) How do these strategies drive the development of conflictual Twitter interactions? (3) What is the primary makeup of linguistic politeness and how is it used to modify the intensity of conflictual Twitter interactions? To investigate these issues, this study integrates data collection and analysis approaches from corpus linguistics, sociolinguistics and quantitative analysis into pragmatic research. It establishes a multidisciplinary research framework that is pragmatics-based and socially oriented. It builds a corpus of Twitter conflictual interactions which will be analysed by using DA and facework, and the findings and user assumptions underlying language use will be verified using two online questionnaire surveys.

4.2 Construction of the Conflictual Twitter Interaction Corpus

This section first introduces the principles of data collection under study. Twitter interactions are organised differently from face-to-face interactions involving specific features that impact how people relate to each other and interpret interactions. In order to determine how Twitter users contribute to discussions and interact with others, Section 4.2.1 discusses in detail the structure of Twitter interactions based on several pilot studies. Section 4.2.2 then presents criteria for classifying Twitter data to ensure that the corpus includes a representative range of Twitter interactions. The second part describes the design of the Conflictual Twitter Interaction Corpus. Section 4.2.3 introduces a three-step process for collecting each genre of conflictual interaction on Twitter. Section 4.2.4 creates an annotation system for the corpus and designs a questionnaire to validate this annotation system.

4.2.1 Structure of Twitter interactions

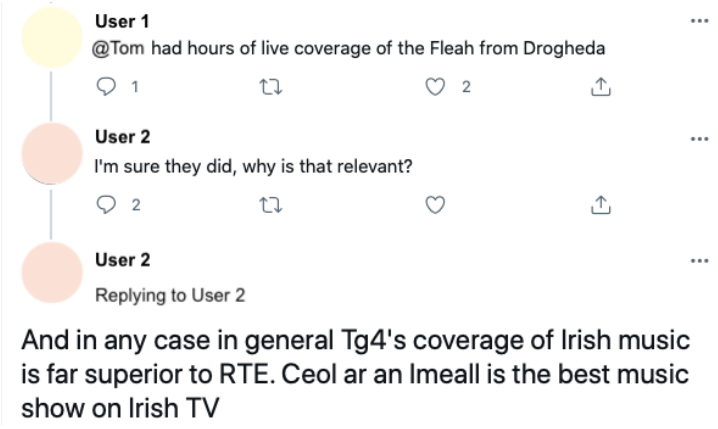
The review found that the present study should collect Twitter interactions that consist of posts with explicit reply-relationships, rather than arbitrary and isolated posts. A case in point is Bou-Franch and Blitvich's (2014) study. In order to build a corpus of conflictual interactions on YouTube, they observed the reply-relationships between scattered comments and rearranged these comments into interactions (see Section 2.1.2 in Chapter 2). According to Scheffler (2017), Twitter users interact with others by commenting on their posts (e.g., tweets, replies), thus developing interpersonal interactions on Twitter that start with the original tweet and display a tree-like structure (see Section 3.1.1 in

Chapter 3). Up to date, however, this tree-like structure has not been investigated in depth and the DA framework has not been well adapted to the structure of Twitter interactions either (see Section 3.3.3 in Chapter 3). In order to obtain more information on these matters and to find ways to address them, I carried out several pilot studies in 2019. I randomly selected a number of hashtags/topics (e.g., #Dublin, #Climatechange) and observed Twitter users' discussions around these topics. Some certain questions received particular attention such as how Twitter users interact through the reply function, how their interpersonal interactions are structured and how their viewpoints are distributed in the interaction.

The result shows that although Twitter users can post only one reply at one time by using the reply function, they may also comment on the previous speaker's post multiple times, resulting in their turn containing more than one reply. For example, in Figure 3, User 2 responds to User 1 by posting two replies, treating User 1 as the intended recipient. The two replies constitute User 2's turn. Hence, the target addressee of the speaker is the Twitter user to whom they are going to response. A single turn on Twitter is not equal to each tweet; instead it refers to all replies published by a user in response to the previous speaker because they together form the complete set of that user's opinions triggered by the previous user's utterance.

Figure 3

Replies and turns

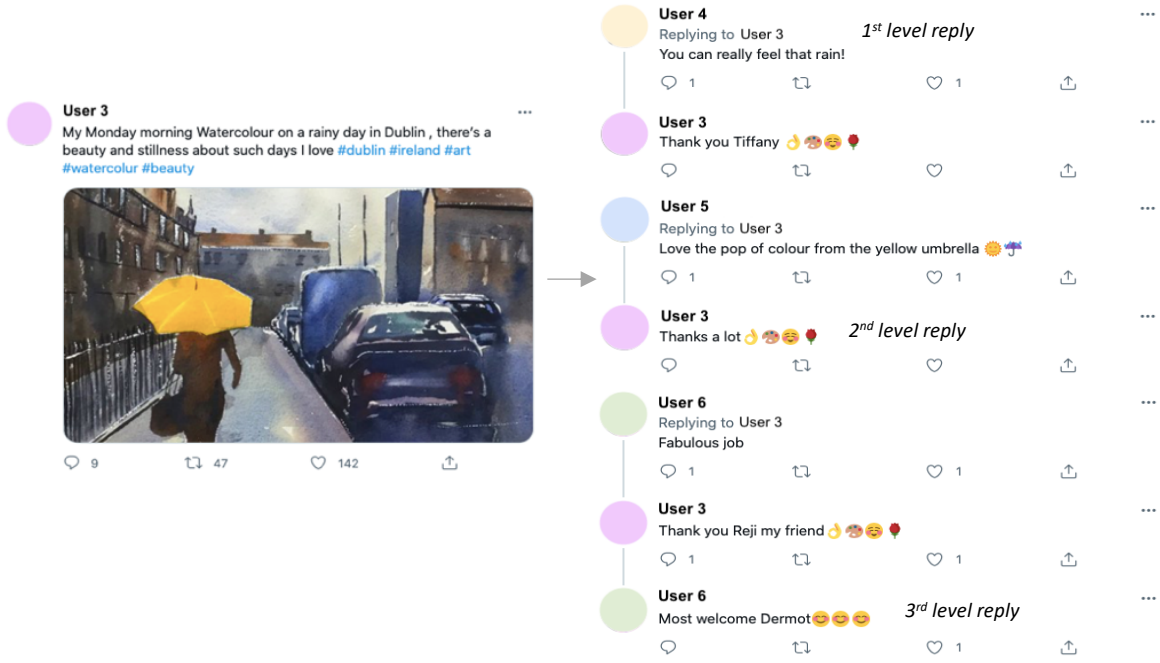


When clicking on a tweet or retweet, replies on Twitter are displayed at different levels. For example, there are seven replies in Figure 4 shown under the original tweet of User 3. Three of them (i.e., the replies of Users 4, 5 and 6) that comment on User 3's original tweet directly are the closest to it. There are also three replies of User 3 placed in the middle, functioning to respond to the replies of Users 4, 5 and 6, respectively. User 6's

comment on one of User 3's replies is farthest from User 3's original tweet. This distribution suggests that replies are organised sequentially on Twitter, and the level of replies is defined in this study as follows: the reply that responds to the original tweet directly is termed as the *1st level reply*; the reply that comments on the 1st level reply is named the *2nd level reply* and so on.

Figure 4

Level of replies

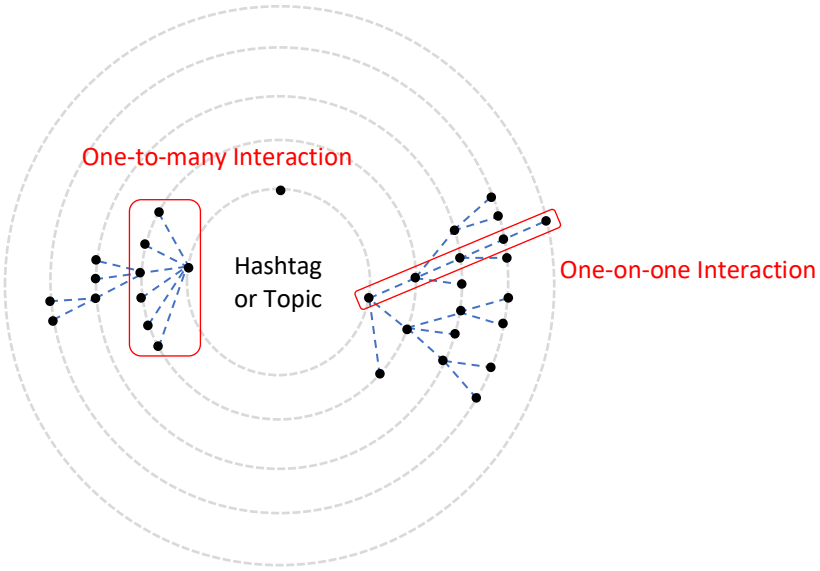


Twitter's reply function allows each level of reply to develop its own discourse. Replies with a mutual-replying relationship are thus picked out and aggregated. Every time Twitter users click on the reply button, a new comment thread is created on Twitter, leading to different types of interpersonal Twitter interactions. For example, in Figure 4, User 3 talks about their feelings about rainy days in Dublin. Users 4, 5 and 6 express their consent and support at the 1st level, User 3 then shows their gratitude at the 2nd level, and User 6 responds to User 3 at the 3rd level. These posts can be viewed from a longitudinal or cross-level perspective as three one-on-one interactions: (1) User 3–User 4–User 3, (2) User 3–User 5–User 3 and (3) User 3–User 6–User 3–User 6. From a latitudinal or same-level perspective, there is: (1) one original tweet from User 3, (2) three 1st level replies from Users 4, 5 and 6, (3) three 2nd level replies from User 3 and (4) one 3rd level reply from User 6. The three 1st level replies establish a one-to-many interaction with User 3's original tweet, but no such interaction is formed between the 2nd level replies and the 1st level replies because the

three 2nd level replies respond to different 1st level replies. This suggests that on Twitter, there are two main types of interpersonal interactions, which can be found in Figure 5: first, *one-on-one interaction* refers to the longitudinal interaction that includes an original tweet and a string of corresponding upper-level replies from different levels. Second, *one-to-many interaction* refers to the latitudinal interaction that consists of a reply or an original tweet with its corresponding upper-level replies at the same level.

Figure 5

One-on-one interactions vs. one-to-many interaction

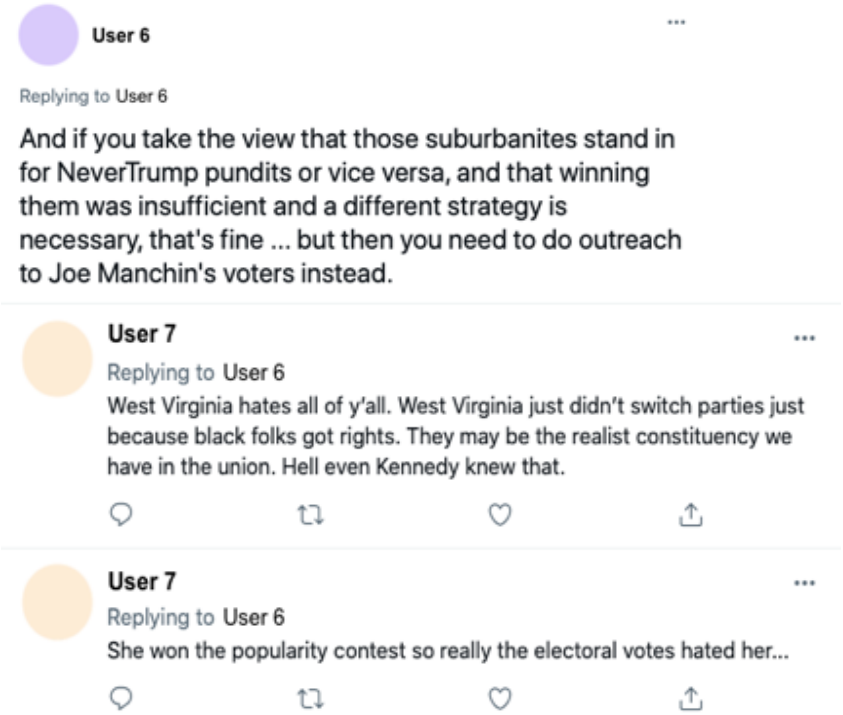


Twitter interactions are very flexible because they can refer to either a comprehensive interaction that consists of a tweet and all posts triggered by that tweet (hereinafter referred to as *CONVERSATION*) or a one-on-one interaction that derives from an original tweet and ends with no more replies (hereinafter referred to as *stream*) or even multiple related one-on-one interactions. For example, there are three streams in Figure 4, which are one-on-one interactions between Users 3 and 4, between Users 3 and 5 and between Users 3 and 6. They are part of the CONVERSATION that derives from User 3's original tweet. In many cases, however, an interaction may also include more than one stream because in order to stay within the word limit of a tweet or to convey a particular meaning by changing the format, Twitter users sometimes split their turn into multiple replies at the same level. For example, Figure 6 shows an excerpt of the interaction between Users 6 and 7. This interaction begins with User 6's tweet, passes through several User 6's self-replies and ends with User 7's reply. In order to respond to User 6's last post that is shown in Figure 6, User 7 publishes two replies at the same level, thus appearing to develop

two streams. The second reply, however, serves to complement the content of the first one, and together they consitute User 7's turn in response to User 6, forming the only interaction in Figure 6.

Figure 6

Streams vs. interactions



This section provides a comprehensive analysis of the structure of Twitter interactions, which constitutes one of the bases for the collection of Twitter data, the differentiation of conflictual from harmonious exchanges of Twitter interactions and the use of the DA framework for their analysis. It is unfeasible for this study to process a large volume of data, due to the impossibility of identifying representative features and structures of conflictual Twitter interactions at the outset. Therefore, it is necessary to develop criteria for sampling conflictual Twitter interactions in order to build a representative corpus that is also suitable for a manual qualitative analysis, which will be discussed in the next section.

4.2.2 Data diversity

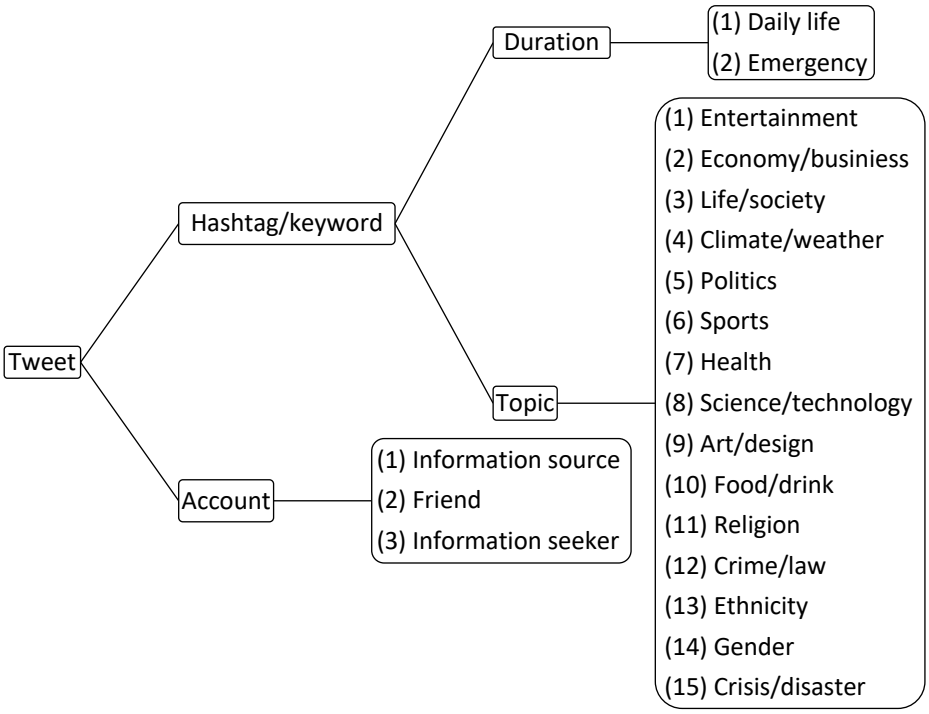
Conflictual interactions in face-to-face and virtual contexts are not homogeneous across contexts, and their distribution has also been found to be socially and contextually variable in the two contexts (e.g., Goodwin & Goodwin, 1987; Johnstone, 1989; Brown, 1990; Herring, 1994, 1996b; Kakavá, 2002; Edstrom, 2004; Williams, 2005; Habib, 2008; Behnam & Niroomand 2011; Koczogh, 2012; Yang, 2013; Decock & Spiessens, 2017ab;

Suroiya, 2017; Bova & Arcidiacono, 2018; Toomaneejinda & Harding, 2018; Fernández-Amaya, 2019; Khammari, 2021). This study assumes that this is also the case for Twitter interactions. In order to properly account for the diversity in conflictual Twitter interactions, it is crucial to ensure that data are drawn from a variety of interactive types or contexts on Twitter.

Shum and Lee's (2013) work provides a solid basis for developing a feasible data collection strategy that will produce a representative corpus (see Section 2.1.1 in Chapter 2). In order to explore a general rule that explains, to the greatest extent, disagreement strategies on two Hong Kong Internet discussion fora, they selected target posts that are both highly controversial and less controversial and collected all replies triggered by these posts without further demographically or socially restricting the participants. Taking this approach, their interactive data includes conflictual exchanges that play a different role in each interaction and thus ensured that these data closely represent the different ways in which users on the two Hong Kong social media sites express disagreements. Conflictual interactions on Twitter similarly derive from an original tweet and are then expanded and developed as users constantly post replies to that tweet. In order to explore how Twitter users generally engage in conflictual interactions, this study should therefore collect a diverse set of Twitter interactions in a coherent manner that also maintains/preserves their real-life complexity. In other words, the interactions in the corpus should cover a wide range of topics that Twitter users from a diverse range of backgrounds are concerned with. As in the case of Shum and Lee (2013), when building a corpus for this study, it is necessary to first of all select diverse types of tweets, which requires this study to establish a classification standard of tweets, as shown in Figure 7. The process for setting up this standard is as follows:

Figure 7

Classification standard of tweets



In order to develop the parameters for classifying tweets, the concept of *speech community* was first considered because it is a basic, however much contested unit, in sociolinguistic research on face-to-face interactions (Patrick, 2004). Researchers argue that the members of a community share one or more things in common (e.g., Patrick, 2004; Meyerhoff, 2011/2006; Milroy & Gordon, 2003; Tagg, 2015). For example, people know each other, have similar types of social status, share common purposes, engage in the same practices, believe that community members have similar language use (e.g., pronunciation, shared jargon, slang and in-jokes) and interactional styles that can be used to distinguish them from other communities and to evaluate language practices in similar ways.

Tagg (2015) suggested that there are two types of speech communities in online interactions: (1) thematic groups and (2) social networks, which are characterised by common goals and interpersonal connections, respectively. On Twitter, users usually add hashtags to mark specific topics, events or issues that are related to their tweets in order to achieve information exchange (Bruns & Moe, 2014). Hashtags, a combination of the hash symbol '#' and keywords, divide people into different thematic groups, constituting groups that deal with daily life or emotions and groups that discuss specific events or problems. By tweeting with the same hashtags, users gather scattered tweets on the same topic into a

larger theme and create the basis of Twitter communities featured by common goals (Bruns & Burgess, 2011; Murthy, 2013; Bruns & Moe, 2014; Halavais, 2014; Gleason, 2016; Gruber, 2017). Through tracking hashtags, Twitter users can not only participate in community discussions but also get to know other users with similar points such as interests, experiences and opinions. As for Twitter interactions featured by interpersonal connections, the follow function plays an important role: users can freely follow others whom they are interested in without permissions or follow-backs (Gruzd et al., 2011). As soon as a follower-followee relationship is established, Twitter users are able to read followees' updates and to interact with followees and their own followers by using the reply function.

Based on Tagg's (2015) work, this study, therefore, argues that thematic groups and social networks are also the two predominant types of communities on Twitter that determine the different nature of Twitter interactions. Interactions in thematic groups develop around discussion topics, that is, they are topic-driven interactions; while interactions in social networks are around Twitter accounts, that is, they are account-driven interactions. In order to capture conflictual interactions of both, this study selected 'hashtag/keyword' and 'account' as the main parameters at the first level.

Since Twitter users need to add hashtags (or keywords) in tweets or replies to participate in topic-driven interactions, the classification of the hashtags (or keywords) is useful for further classifying such interactions. Studies on hashtag classification mainly come from the field of computer science and focus on improving the accuracy of the classification by optimising natural language processing techniques (e.g., Lee et al., 2011; Romero et al., 2011; Cui et al., 2012; Quercia et al., 2012; Ma et al., 2012; Posch et al., 2013; Bogdanov et al., 2013; Jeon et al., 2014; Rosa et al., 2014; Li et al., 2016; Iman et al., 2017; Nurfikri & Mubarok, 2018; Tiwana et al, 2018; Yu et al., 2019). Although the results of the classification are not their focus, they do suggest two critical classification parameters: 'duration' and 'topic'.

In terms of duration, some topics last for a long period of time and are generally about daily life, activities and thoughts, but others are usually short-lived topics about emergencies. This classification categorises hashtags explicitly according to the time information provided by Twitter and is well accepted by researchers (e.g., Cui et al., 2012;

Liu et al., 2016; Visansirikul & Kitisin, 2018; Tiwana et al, 2018). 'Daily activities' and 'emergencies' thus become the sub-parameters of the parameter 'duration'.

As for topic, although the classification of hashtags (or keywords) often varies from person to person, this does not pose a big problem when selecting the target tweets. This is because a tweet is often related to several aspects of social life and the diversity of interactions would not be significantly diminished due to the different classifications of the topics of the hashtags/keywords in target tweets. For example, a tweet that promotes a vegetarian restaurant is not only about food or drink but also about a kind of lifestyle or even about a vegan's self-identification. This study thus identified the sub-parameters of the parameter 'topic' by summarising what are commonly used in studies on hashtag classification (e.g., Romero et al., 2011; Lee et al., 2011; Ma et al., 2012; Cui et al., 2012; Bogdanov et al., 2013; Li et al, 2016; Iman et al., 2017; Nurfikri & Mubarok, 2018; Yu et al, 2019). As a result, a total of 15 parameters were selected, including: 'entertainment', 'economy/business', 'life/society', 'climate/weather', 'politics', 'sports', 'health', 'science/technology', 'art/design', 'food/drink', 'religion', 'crime/law', 'ethnicity', 'gender' and 'crisis/disaster'.

The most basic and common measure of account-driven interactions is presumably the 'follower-followee ratio' (e.g., Yan et al., 2018; Saito & Masuda, 2013; Anger & Kittl, 2011). The *follower-followee ratio* compares the number of users following a certain Twitter account with that of users being followed by that Twitter account. According to Yan et al., (2018), Twitter accounts are termed as *information sources*, if the number of their followers is greater than that of their followees; Twitter accounts are called *friends*, if the number of their followers is similar to that of their followees; Twitter accounts are named as *information seekers*, if the number of their followers is much lower than that of their followees. This study thus selected 'information seeker', 'information source' and 'friend' as the sub-parameters of the parameter 'account'. In this way, conflictual interactions can be collected from a variety of interactive types or contexts on Twitter and can be disseminated to a broader audience, rather than being restricted to those who follow large media outlets. The step-by-step collection procedure is discussed in detail in the next section.

4.2.3 Collecting conflictual Twitter interactions

Data collection began in July 2020 over a two-month period and involved three steps, as can be seen in Figure 8. First of all, this study selected and collected suitable target tweets, whose amount is appropriate for a manual qualitative analysis, according to the hashtags (or keywords) and Twitter accounts specified in Section 4.2.2 (see Figure 7). The second step is to extend the collected tweets into interpersonal Twitter interactions by systematically collecting posts triggered by these tweets based on the observation of the structure of Twitter interactions in Section 4.2.1. Since there are both harmonious and conflictual exchanges in Twitter interactions, the final step is to distinguish them based on the observation of opinion linkage and to extract conflictual Twitter interactions. By following these three steps, it is possible to develop a corpus that is suitable for investigating conflictual interactions on Twitter. The specific procedures are as follows:

Figure 8

Steps for collecting conflictual Twitter interactions

Step 1: selecting target tweets

In order to select suitable tweets, three additional factors have to be integrated with the parameters decided in Section 4.2.2 (see Figure 7). First, the tweets should have a relatively high number of retweets and replies because such tweets are more likely to generate a sufficient variety of Twitter interactions. Second, the tweets should be written in English. Since the review found that Brown and Levinson's (1987/1978) facework is more suitable for analysing interactions in English-speaking cultures (see Section 3.4 in Chapter 3), this study needs to select English-written tweets to weaken the controversies arising from the cultural adaptability of this theory. Although such tweets are not necessarily posted by English-native speakers, their authors are at least able to express their opinions in English fluently and have knowledge of English etiquette. While code-switching may occur in

interactions originating from English-written tweets, posts that include other languages are also retained because the non-English parts may serve as a strategy to intensify or mitigate conflict in English-based interactions. Third, to avoid English-written tweets being geographically concentrated, their authors should come from countries where English is not widely used as a first or primary language and countries where English is widely used as the main or only language of communication.

To be more specific, this study randomly and broadly picked about 50 keywords/hashtags according to the classification standard shown in Figure 7. After searching for each of these keywords/hashtags on Twitter, it subsequently chose 50 English tweets that had triggered heated debate on Twitter and was displayed at the top of the research results. Lastly, these tweets were verified with their organisers' properties and were sorted by the account classification standard shown in Figure 7. Using the process, this study finally selected 38 tweets (including 36 original tweets and two retweets). From the perspective of topic, the 38 tweets covered the 15 categories mentioned above, including entertainment (e.g., Game of Thrones), economy/business (e.g., redundancy), life/society (e.g., vegetarian lifestyles), climate/weather (e.g., climate strike), politics (e.g., election), sports (e.g., American football), health (e.g., Covid 19), science/technology (e.g., 5G), art/design (e.g., iPhone), food/drink (e.g., gold chicken wings), religion (e.g., Christians), crime/law (e.g., terror attacks), ethnicity (e.g., the death of Floyd), gender (e.g., drag queen), crisis/disaster (e.g., Hurricane Dorian). In terms of duration, 28 tweets are about daily life, and 11 tweets are about emergencies including Notre Dame fire, Amazon fire, Hurricane Dorian, Climate Strike, Australian bushfire, the death of Soleimani, the death of Kobe, the death of Floyd, Covid-19, Beirut blast and the Wall of Mom's protest. From the aspect of account, 26 of the 38 tweets are from 'information source' accounts, 10 tweets are from 'friend' accounts, and three tweets are from 'information seeker' accounts. This study, therefore, ensured that the interactions collected based on these 38 tweets are efficiently diverse regarding content.

Step 2: collecting Twitter interactions

Step 2 discusses the procedure of how this study collected Twitter interactions based on the 38 tweets decided in the first step. As the 38 tweets have diverse origins, as long as the interactions derived from the 38 tweets are also varied in form, a diverse set of Twitter

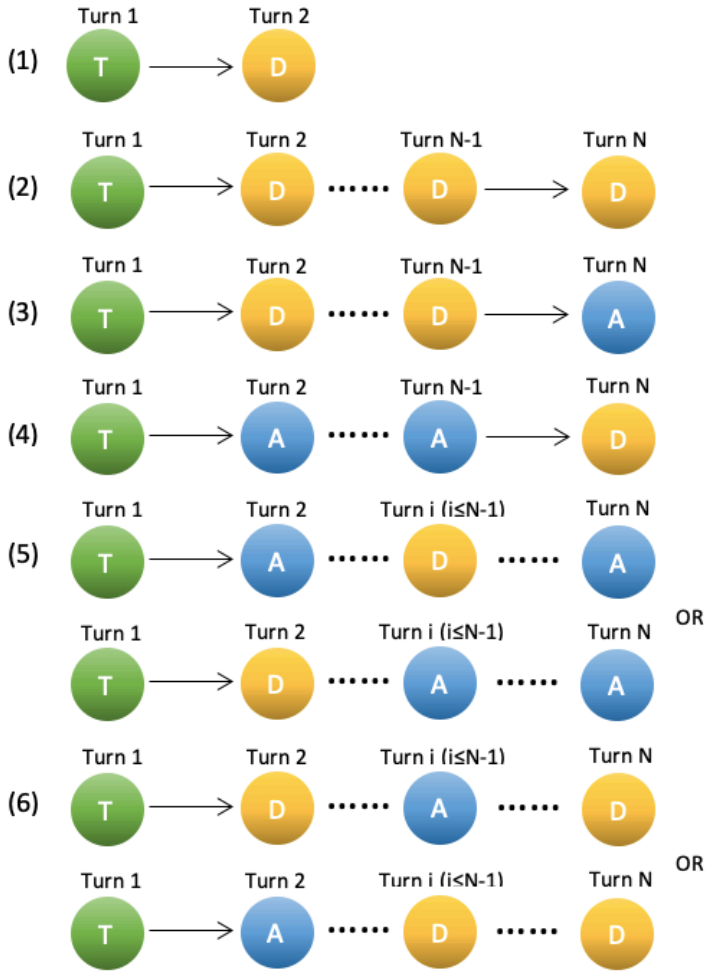
interactions can be collected. The analysis of the structure of Twitter interactions in Section 4.2.1 suggests that there are two types of interpersonal interactions on Twitter, namely one-on-one interactions and one-to-many interactions (see Figure 5). Hence the corpus should include both types of interactions. Since any reply or tweet can be retweeted or commented on an unlimited number of times, the interactions triggered by a tweet can literally be prolonged indefinitely. The number of posts in the corpus, however, cannot be too large because of the feasibility of a manually qualitative analysis. The posts in the corpus also have to comply with the constraint that the initial tweet of an interactive stream in which it is located should belong to the 38 tweets. Based on these criteria, this study first collected 38 tweets, all replies triggered by these tweets (which were published by September 2020) and the reply-relationships between these posts. This resulted in 38 Twitter CONVERSATIONs with a total of approximately 34,000 words, including the content of posts, Twitter users' screen names and the time of publication. The data was stored in Word documents. The multimodal features and special formats in posts were transcribed into text: (1) pictures, GIFs and videos were encoded as '[PIC/GIF/VIDEO: Link]'; (2) emoji were encoded as '[the emoji name]' (e.g., [Red Heart]); (3) abnormal white spaces were changed into '[NWS]', indicating the number of abnormal white spaces; (4) at the beginning of each row, row numbers were marked as '[NstR]'; and (5) empty rows were noted as '[NstR: ER]'.

Step 3: extracting conflictual Twitter interactions

An initial analysis indicates that the majority of interactions contained in the 38 Twitter CONVERSATIONs are conflictual and that there are the six types of relationships between adjacent turns shown in Figure 9. In a conflictual interaction consisting of two turns, the second turn disagrees with the target opinion. If there are N (N>2) turns in a conflictual interaction, there may be N-1 disagreements followed by the target opinion, or N-2 disagreements between the target opinion and the agreement in the last turn, or N-2 agreements between the target opinion and the disagreement in the last turn. If the N-2 turns in an interaction between the target opinion and the agreement/disagreement in the last turn include both agreements and disagreements, this interaction is also conflictual. In the 38 Twitter CONVERSATIONs, there are also a handful of harmonious interactions (with fewer than 2,000 words), in which all users hold the same/similar opinion to the original tweet.

Figure 9

Opinions linkages in conflictual interactions



This study removed these harmonious interactions from the collected 38 Twitter CONVERSATIONs for two reasons. First, they are beyond the research scope of the present study because the nature of conflictual interactions is controversial (Arcidiacono & Pontecorvo, 2009). Second, as the analysis of opinion linkages in Figure 9 shows, turns in conflictual interactions are not always mutually refuting and such interactions often include several successive turns that share the same or similar perspectives. The Conflictual Twitter Interaction Corpus was then established by extracting conflictual interactions from the 38 Twitter CONVERSATIONs. In this corpus, there are a total of 11,226 posts (including tweets and replies) and 5532 streams with approximately 32,000 words. Due to the blurry and flexible nature of the boundary of Twitter interactions, as discussed in Section 4.2.1, the amount of conflictual interactions in this study was not counted. There is no need to separate CONVERSATIONs into sub-interactions for counting purposes, but researchers are able to set a specific boundary according to their research needs. To facilitate a detailed

analysis of the corpus, this study also annotated the data in the corpus and the specific rules are presented in the next section.

4.2.4 Annotating conflictual linguistic strategies

The review of corpora in Section 3.1.2 in Chapter 3 found that most ready-made annotation tools mainly achieve a reasonable degree of accuracy in automatically detecting and annotating various kinds of grammatical, lexical and syntactic information. When used to detect the meaning of language in interactions, these tools are far less accurate (McEnery et al., 2006). This study, therefore, should create a set of rules for annotating the corpus manually. Linguistic research on the structure of conflictual interactions also suggests that conflictual interactions can be roughly divided into two stages based on changes in intensity: (1) escalation and (2) de-escalation (see Section 2.2 in Chapter 2). This study set out to annotate the linguistic strategies employed by Twitter users in the corpus to escalate or to de-escalate conflict. However, the analysis did not identify the de-escalation stage in the corpus. The possible reasons for this absence are discussed in Chapter 5 (see Section 5.1.2; see more discussions in Section 7.3 in Chapter 7).

In terms of escalation strategies, current studies on conflictual interactions suggest that there are two possible types (see Section 2.1 in Chapter 2): (1) disagreements and (2) personal attacks. This study thus selected three tags: (1) 'disagreement', (2) 'personal attack' and (3) 'other' and created a set of annotation rules according to existing definitions. That is, when tweeting (e.g., texting, quoting, inserting emoji/emoticon/GIF/video, etc.):

(1) if a Twitter user refutes their opponents' viewpoints, this user's post will be marked as '*disagreement'*;
(2) if a Twitter user's contribution is directed against their opponents rather than the position their opponents maintain, this user's post will be marked as '*personal attack'*;
(3) if a Twitter user performs actions other than the two mentioned above, this user's post will be marked as '*other'*.

For example, in Excerpt 8, in order to illustrate that User A1's opinion is "completely and utterly invalid" (Line 09), User B1 points out that User A1 is a minor, a centrist and a furry. User B1 obviously attacks User A1 and User B1's reply was thus tagged as 'personal attack'.

Excerpt 8 (Interaction 1)

Line	User Account	Text	Level	Turn
05	User A1	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
06				
07				
08	User B1	[1R]you're 15, anti-ACAB, a centrist, and a furry. your opinion is completely and utterly invalid	2nd	5
09				

In Excerpt 9, Users A2 and B2 pray for the people who are suffering due to Hurricane Dorian, which is opposed by User C2. User C2 does not attack Users A2 and B2 but explains the reasons for their objection. They imply that Users A2 and B2's prayers are useless to the victims by arguing that God is not omnipotent. User C2's reply was thus marked as 'disagreement'.

Excerpt 9 (Interaction 2)

Line	User Account	Text	Level	Turn
05	User A2	[1R]Father in Heaven, please safeguard Your people in the path of Hurricane Dorian. Loose Your mighty angels to minister unto each of them at this time. Fill them with Your peace that passes all understanding. Keep them together in love. In Jesus' Name, Amen.	1st	2
06				
07				
08				
09				
10	User B2	[1R]And the animals and wildlife too, Lord, created by you.[Folded hands][Folded hands][Folded hands]	2nd	3
11				
12	User C2	[1R]I'm assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn't sound like a very compassionate deity to me.	3rd	4
13				
14				
15				

Excerpt 10 shows an interaction between Users C3 and D3. User C3 challenges the criticism from User D3 in the previous turn (which was not shown in Excerpt 10). In response, User D3 explains their reason for taking User C3 as thin-skinned, "because you think being called a loser is abuse" (Line 29), thus attacking User C3 for being a loser. User D3 then discusses their reasons for objecting to User C3's opinion in a second reply and points out the issue of User C3's examples in the previous discussions, "your one example is a misquote that got fixed to mean the opposite of what you wanted it to mean" (Line 31–33). Hence User D3's turn (including two replies) was labelled as 'personal attack' and 'disagreement'. Note that in the annotation, the key to distinguishing personal attacks from disagreements is to discern what the speaker is seeking to oppose (i.e., the speaker's communicative intention is directed towards the recipient's opinion or the recipient themself), rather than the means employed by the speaker, such as the recipient's body, health, reputation or family members. The reason for this is that it is possible for the speaker to attack the recipient's personal characteristics when arguing against the recipient's views.

Excerpt 10 (Interaction 3)

Line	User Account	Text	Level	Turn
28	User C3	[1R]Why thin skinned?	9th	10
29	User D3	[1R]Because you think being called a loser is abuse[Face with tears of joy]	10th	11_1
30				
31	User D3	[1R]It doesn't change the fact that your one example is a misquote that got fixed to mean the opposite of what you wanted it to mean	11th	11_2
32				
33				

Researchers' annotations are very likely to be influenced by their cultural background and other interpersonal factors, leading to researcher bias, which is unavoidable but can be minimised. In order to ensure that these annotation rules are not idiosyncratic to the researcher, this study also conducted a survey by using the aforementioned examples, namely Excerpts 8, 9 and 10. The questionnaire is displayed in Appendix 2 (see Questionnaire 1). It distributed 15 questionnaires to people over the age of 18 including those with and without linguistic training and asked them to invite their adult friends to participate in the survey. A total of 30 returned questionnaires were finally collected. Before filling in the questionnaire, participants were required to sign a consent form. In the survey, they were given the three interactions shown in Excerpt 8, 9 and 10 and were asked to annotate the replies of Users B1, C2 and D3 following a three-step training. Participants were first asked to explain what Users B1, C2 and D3 are communicating to their opponents. They were then asked to select which actions Users B1, C2 and D3 perform against their opponents among the given selections (i.e., disagreement, personal attack, other). Finally, they were provided with the annotation rules and were asked to annotate the replies of B1, C2 and D3.

The result shows that 83% (25) of the participants marked User B1 the same as I did, an increase of 19% (4) over the pre-training result. Although only 15 participants shared my opinion that User C2 was disagreeing with the opponent, almost all of them who selected 'other' (3) believed that User C2 was expressing an opinion. For example, some argued that User C2 was "sharing an opinion" and others believed that User C2 was stating "another side of the problem". Both of these figures were consistent before and after training. In terms of User D3's reply, since there was no hint in the question that this reply could be annotated with multiple strategies, only 10% (3) of the participants labelled this reply as a combination of 'disagreement' and 'personal attack'. However, the number of participants who marked it as a disagreement and those who annotated it as an attack were both significant, around 63% (19) and 27% (8), and the number of participants who labelled it as a combination

showed an increase of one participant as compared to the pre-training result. These results suggest that the annotation rules used in this study are efficient and reasonable and were subsequently used to annotate every turn in the Conflictual Twitter Interaction Corpus. The next section discusses suitable approaches for analysing the corpus in detail.

4.3 Data analysis

This section describes the methods used in the present study to analyse how Twitter users employ conflictual linguistic strategies to engage in conflictual interactions. Overall, this study takes a combined qualitative and quantitative approach. The three analysis chapters address each of the three research questions, respectively. Chapter 6 analyses the escalation strategies in the corpus and preliminarily examines their linguistic nature and their interactional functions. Chapter 7 provides an in-depth analysis of the role of personal attacks in the development of conflictual Twitter interactions. Chapter 8 explores the makeup of linguistic politeness and its function in adjusting the intensity of conflictual interactions. The specific analytic methods in the analysis chapters are as follows:

The textual analyses in Chapters 5, 6 and 7 largely rely on combining DA with Brown and Levinson's (1987/1978) facework. This combination allows this study to examine how speakers decide the degree of damage to their opponents' faces and to further analyse how the construction of the speaker's turn is influenced by the previous turn and how it in turn affects the design of the next turn. Taking this approach, conflictual Twitter interactions can be explored from a dynamic perspective, thus revealing the linguistic mechanisms of their development. This analytical approach avoids a one-sided understanding of the function of linguistic strategies due to an excessive focus on the speaker's ability or intention in language use and prevents neglecting the design of each turn because of an overemphasis on the overall structure of conflictual Twitter interactions. For example, Excerpt 11 shows a snippet of an interaction between Users B4 and C4 about whether Exxon intentionally conceals the damage of fossil fuels to climate change, which was analysed in the pilot studies. In the previous turn (not shown in Excerpt 11), a Twitter user (not Users B4 or C4) provides evidence of Exxon's denial of climate change. They share a screen recording which shows that a video is inserted in the middle of an article on the website of *The New York Times.* This article documents a court case where Exxon was charged with lying about the costs and consequences of climate change. The inserted video, however, makes propaganda

about how eco-friendly the Exxon's industries are. The previous speaker uses this screen recording to prove that until now Exxon is still misleading the public about climate change. At the 2nd level, User B4 entered the discussion and started a one-on-one interaction with User C4. User B4 presents their feelings about the video recording, "that's odd" (Line 08), and asserts, in line with the previous speaker, that Exxon is dodging the truth because all industries affect nature (Lines 09–11). User C4 (Lines 12–16) openly voices a contrary perspective. However, instead of directly confronting User B4, they couch it as a request for clarification, "I'm not sure I follow the point […] who could know for sure?" (Lines 12–16), which shows User C4's protection of User B4's negative face. In response, User B4 still treats User C4's response as a face-threat and refuses to confirm or engage with it. They replicate User C4's escalation approach by requesting User C4 to watch the video before asking questions, "but perhaps you listen to the interrogation of the scientist who worked for #exxonknew, if you have any questions" (Lines 17–19). Given User C4's protection of User B4's face, User B4, however, also mitigates the threat to User C4's negative face. They introduce the request with an apology, "sorry" (Line 17), and express it with a suggestion marker, "perhaps" (Line 17).

Excerpt 11 (Information 4)

Line	User Account	Text	Level	Turn
08 09 10 11	User B4	[1R]that's odd. to say it clearly, #ExxonKnew[1WS], no, industry is NOT able to do what nature is doing, that is the problem. Nature regenerates and creates, industry destroys. end of cake. we are not God. [2R:ER] [3R]2	2nd	3
12 13 14 15 16	User C4	[1R]I'm not sure I follow the point [2R:ER] [3R]1) #ExxonKnew at the very least that there was a very high risk of bad things happening, but they pushed the narrative that there was none [4R]2) Bad things have started happening [5R]3) It's not their fault at all because who could know for sure?	3rd	4
17 18 19 20 21 22 23 24	User B4	[1R]Sorry, but perhaps you listen to the interrogation of the scientist who worked for #exxonknew[1WS], if you have any questions , that they knew exactly what a)they were doing b) was comming c) that they are guilty. [2R]https://twitter.com/YEARSofLIVING/status/118748872668 15072256 [3R][Quote: https://twitter.com/Independent/status/11877192065629102 09]	4th	5

As Excerpt 11 suggested, this study first observed Twitter users' language choices and expression patterns across a large number of interactions. Through using Brown and Levinson's (1987/1978) facework, it then investigated how Twitter users protect or hurt their opponents' faces, how they react to their opponents' damage to their own face and how

they express their opinions about people and issues across a number of different hashtags/topics. Note this analysis does not aim to determine or verify exactly how Twitter users felt or why their actions were performed in the manner displayed in interactions. To conduct facework, researchers must select the most appropriate possibilities for creating a closed logical circle of interlocutors and thus formulate a rational and reasonable explanation from a professional's perspective. Identifying individuals' innermost motivation and feelings falls outside the scope of facework. This study acknowledges, however, that researcher bias is present in the analysis. In order to reduce the impact of researcher bias on research results, this study ultimately examined Twitter users' language from a dynamic perspective by using DA, with a special focus on the recipient's interpretation. It analysed how Twitter users construct their turns and what this suggests about their interpretations of the previous turn and also surveyed how their turns affect the next turn. That is, it examined how the users' turns are interpreted by the next speaker and how the next speaker designs the next turn based on their understanding.

A quantitative analysis of the frequency proportion was also adopted in Chapters 6 and 7 to demonstrate the feasibility and significance of an in-depth analysis of personal attacks and linguistic politeness. The *frequency proportion* is used to show how common a particular construction is (Warne, 2018). The formula is: *Proportion = f/n.* In the formula, *f* refers to "the frequency of a response" and *n* is "the sample size" (Warne, 2018, p. 37). Pearson's *r* was also employed based on the Excel function 'CORREL' in Chapter 6 to analyse whether the prevalence of personal attacks is related to contextual features (e.g., technique issues, the nature of interactions). The aim is to get a better understanding of the distribution of attacks. Pearson's *r*, is one of the most common correlation coefficients in social science, ranging from -1 to +1 (Warne, 2018). The further *r* is away from 0, the stronger the relationship between the variables; conversely, it suggests a weaker relationship between the variables. A positive *r* indicates a positive correlation between the two variables, meaning that the score of the dependent variable increases (or decreases) as the score of the independent variable increases (or decreases). A negative *r* means that there is a negative correlation between the two variables, meaning the score of the dependent variable increases (or decreases) as the score of the independent variable decreases (or increases). However, when *r*=0, there is no relationship between the variables.

In other words, the scores of the independent variable do not form any pattern with the scores of the dependent variable.

In addition, in order to verify findings about the role of linguistic politeness in escalation strategies and to obtain more information that is not easily available from the textual analysis alone, a questionnaire about Twitter users' attitudes towards language use in conflictual interactions was conducted, the results of which are presented in Chapter 7. The questionnaire is displayed in Appendix 2 (see Questionnaire 2). In this questionnaire, participants who are Twitter users over the age of 18 and are able to communicate fluently in English were recruited in three ways: (1) inviting Twitter users by sending Twitter messages, (2) inviting strangers on the street and (3) inviting colleagues and friends who the researcher knows. Since one of the questions in the questionnaire dealt with people's opinions about feminism, the number of male participants and female participants was also balanced in order to prevent results being influenced by the respondents' gender. A total of 33 completed questionnaires and consent forms were obtained which represents a response rate of about 60%. However, three questionnaires were removed because participants stated that they were not Twitter users.

In the questionnaire, participants were first asked to express their favourite and least favourite aspects about using Twitter to interact. They were then required to read the fragment of one conversation analysed in Chapter 7 (see Excerpts 121 and 122), shown here as Excerpt 12. @Jessica invites women to participate in the discussion, "What would you do if all men had a 9 p.m. curfew?" (Lines 01–02), and repeatedly emphasises that this discussion is for feminists and the topic is about male privilege. User A32 publicly challenges @Jessica's tweet, arguing that @Jessica deliberately provokes women's hatred of men and overtly attacks @Jessica for being "condescending and shallow" (Line 07). However, they also avoid rejecting @Jessica's claim by using rhetorical questions and diminishing the coercion of their request to @Jessica through a "how about" type question (Line 07). In the third task, the participants were required to explain how they designed their replies in the previous task. Since rhetorical questions and suggestions are taken as two manifestations of linguistic politeness in Chapter 7, in order to verify this finding the participants were asked to explain whether they felt offended, whether they tried to be polite to User A32 and whether User A32 was trying to be polite. Finally, the participants were asked to recall their experience of using the sentence "Could you please fuck yourself" in conversations both on

Twitter and in face-to-face contexts. This is a typical phrase identified in Chapter 7, in which the speaker uses polite and offensive expressions. The main purpose of this task was to assess people's understanding of the phrase and their reasons for using it. The results of this survey are presented in detail in Chapter 7.

Excerpt 12 (Interaction 32)

Line	User Account	Text	Level	Turn
01 02 03 04	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #manprivilege #privilege	tweet	1
05 06 07 08 09	User A32	[1R]You do realize the vast majority of dudes don't bother or harrass women right? Your post comes off as incredibly condescending and shallow. How about instead of generalizing you start encouraging women and tell them how to protect themselves if they would like the option?	1st	2

4.4 Ethical consideration

The last section in this chapter discusses the ethical issues involved in collecting and analysing Twitter data and also the solutions to these issues. One of the main issues that this study encounters is whether researchers could use a Twitter user's posts without obtaining their permission. This study first consulted this question in previous studies and found that since Twitter users can decide whether they want their tweets to be public (i.e., all Twitter users could see the messages) or private (i.e., only followers are able to see the messages), using posts on Twitter without their authors' expressed permission is permissible (e.g., Honeycutt & Herring, 2009; Boyd et al., 2010; Murthy, 2013; Bruns & Moe, 2014). The audience of public tweets includes both their followers and Twitter users who do not follow the authors of such tweets, but private tweets are only available to their followers. However, any Twitter user is literally able to become the follower of the author of a private tweet and view this tweet because they can only determine who to follow but not their followers. In addition, since followers can still share screenshots of private tweets and other users can further retweet these screenshots, private tweets are still highly likely to spread widely on Twitter. These facts greatly blur the line between being public and private.

According to the *Twitter Privacy Policy*, Twitter users acknowledge that "Twitter is public and Tweets are immediately viewable and searchable by anyone around the world"

("Twitter—Twitter privacy policy", 2022).[20] They admit that most information on Twitter is public, including their profile information, their time zone and language, the content of their Tweets, the application and version of Twitter they use, the people they follow and who follow them and the tweets they like or retweet. They also agree to grant the Twitter company a worldwide, non-exclusive, royalty-free license (with the right to sublicense) to use, copy, reproduce, process, adapt, modify, publish, transmit, display and distribute such content in any and all media or distribution methods now known or later developed (for clarity, these rights include, for example, curating, transforming and translating) ("Twitter—User Agreement", 2022). The Twitter company also reminds Twitter users that they should think carefully about what they post publicly ("Twitter—Twitter privacy policy", 2021) and provides options in the 'Privacy and Safety' settings to decide whether their tweets are publicly available ("Twitter—Twitter privacy policy", 2022). In the *Developer Agreement*, Twitter grants researchers "a non-exclusive, royalty free, non-transferable, non-sublicensable, revocable license" to "copy a reasonable amount of and display the Twitter content" and to "modify Twitter Content only to format it for display" ("Twitter—Developer terms", 2022).[21] Tweets that can be accessed by researchers who use a Twitter account that does not follow any other users are therefore literally public data.

The 38 tweets collected to build the corpus for this study are all publicly available. To obey Twitter's *Privacy Policy*, collected posts were also removed from the dataset if they were deleted or if their authors' accounts were suspended, withheld under local law or permanently deleted, during the data collection. This study only collected the information that is automatically displayed on Twitter (e.g., user IDs, tweets, replies and specific time), rather than recruiting participants to obtain the collected interactions or tracking the history of their language use on the Internet. Most of the posts were only stored in the corpus to do DA and were not published or republished. This study only provided the direct quotation of interactional fragments as examples, with only a few turns that serve to illustrate the argument. The interactions presented in the Appendix 1 are only the quoted fragments in the paper, not entire interactions. This study also removed and replaced most identifiable information from the quoted fragments, such as users' screen names. Only the screen name of official institutional account was retained.

[20] Accessed, 22nd February 2022: https://twitter.com/en/privacy
[21] Accessed, 22nd February 2022: https://developer.twitter.com/en/developer-terms/agreement

In the following three chapters, this study presents the results of the analysis of the corpus (that has been annotated according to the standard testified by the first questionnaire) and the second questionnaire. It discusses the specific ways in which the interactivity of Twitter conflict is manifested and the particular linguistic mechanisms that contribute to its interactivity based on these results. It also compares these results with the findings in current research on offline conflictual interactions to examine the function and impact of technology in human communication, beliefs about language and language use. Specifically, in order to answer the first question (i.e., What are the linguistic strategies employed by Twitter users to engage in conflictual interactions?), Chapter 5 provides an overview of the conflictual linguistic strategies observed in the corpus. It also examines how Twitter users use these strategies to cooperate with others and to progress the interaction, thus addressing the second question (i.e., How do conflictual linguistic strategies drive the development of conflictual Twitter interactions?) from a relatively theoretical and general perspective. Based on the findings in Chapter 5, Chapter 6 explores the second question in more detail. It provides an in-depth analysis of one conflictual linguistic strategy 'personal attacks'. It quantifies its occurrence in the corpus and investigates its discursive functions in a qualitative manner. Finally, Chapter 7 is devoted to the last question (i.e., What is the primary makeup of linguistic politeness in conflictual Twitter interactions and how does it modify the intensity of such interactions?). It revises Brown and Levinson's (1987/1978) framework based on the corpus, redefines the notion of linguistic politeness for Twitter contexts, investigates each type of linguistic politeness approach in detail paying close attention to its makeup, communicative intent and discursive function. It also discusses the results of the second questionnaire.

Chapter 5 An Overview of Conflictual Linguistic Strategies

5.0 Introduction

As one remarkable feature of CMC (Reid, 1991), deindividuation has led to the prevalence of uninhibited linguistic behaviour on the Internet such as verbal aggression and blunt disclosure because users' behaviour often suspends the social norms that are widely accepted in face-to-face interactions (Murthy, 2013; Haugh & Sinkeviciute, 2019; Graham, 2019; Sifianou, 2019). Twitter, a public-oriented social medium enacting extensive interactions between strangers, has repeatedly been singled out as being particularly riddled with conflictual interactions (Murthy, 2013). However, despite this reputation and its public attention, linguists still know very little about the nature of conflictual interactions in CMC in general and on Twitter in particular. Although conflict on Twitter, like offline conflict, generally has deep socio-historical roots such as political and economic imbalances between states, historical discrimination or prejudice against ethnic or social groups, it is most likely to arise when users perceive themselves as opposing others on the same matter. This conflict also takes place verbally as in offline interactions and is heavily intertwined with users' discussions and debates. It is, accordingly, exacerbated or diminished as the debates evolve and is shown as a chain of linguistic actions in a chronological sequence, which are dominated by uninhibited ones. This discursive nature of conflict leads to an interactive perspective of conflictual Twitter interactions. Exploring the linguistic mechanisms that give rise to online conflict in action under the constraints and shaping of technology thus becomes the major aim of this study. To put the purpose another way: this study aims to explore how Twitter users employ linguistic strategies to do conflictual interactions and how Twitter algorithms impact on users' language uses and interactional patterns (see Section 1.1 & 1.4 in Chapter 1).

In order to respond comprehensively to the first purpose, this chapter provides, for the first time, an overview of conflictual linguistic strategies found on Twitter based on the purpose-built Conflictual Twitter Interaction Corpus. It identifies the types of linguistic strategies used by Twitter users to do conflict and investigates the linguistic nature of these strategies. It also generally discusses the roles of these linguistic strategies in the development of conflictual Twitter interactions by examining the conversational rules that underlie and constrain people's use of these strategies.

Section 5.1 discusses the differences between Twitter and offline conflictual interactions regarding the overall makeup and stage, arguing that conflictual Twitter interactions typically lack the de-escalation stage and that this absence appears to suggest that technology is changing the conversational norms that people generally obey in face-to-face contexts about what is permissible and what is forbidden (see Section 5.1.1). It then discusses the linguistic strategies for escalating conflictual Twitter interactions found in the corpus, including disagreements and personal attacks, and provides a close-up illustration of these two escalation strategies based on examples from the corpus (see Section 5.1.2).

Section 5.2 first explores how Twitter users cooperate in conflictual interactions by using linguistic strategies, thus challenging the stereotype in interactional studies that only agreements can result in cooperation (see Section 5.2.1). The analysis shows that the conversational mechanism of face-to-face interactions identified by discourse analysts (i.e., in the course of a conversation, interlocutors should take turns being the next speaker and issue the next turn) applies to conflictual Twitter interactions as well. It finds that to do conflict, users first become the recipient of the (previous) speaker's turn, interpreting the speaker's turn design and deciding their own strategic goals. Specifically, they need to identify the strategies employed by the speaker, speculate about the speaker's communicative intent, decide whether or not and how to respond to the speaker based on their communicative intent. The users then switch their roles to that of the (next) speaker, designing their own turn based on their interpretations and strategic goals, posting the next turn and also leaving it open for interpretations by their recipients. However, an in-depth analysis in Section 5.2.2 suggests that in this cooperation, Twitter users who assume the role of the recipient may not always understand the use of strategies or strategic goals of Twitter users who act as the (previous) speaker; even if they do, they may not necessarily design their own turn in line with their interpretations. This section thus continues to compare Twitter users' turn designs with those of their recipients in conflictual interactions (see Section 5.2.3). It argues that both types reflect their speakers' interpretations of the previous turn (including its use of strategies and its purposes for using these strategies), as well as their own strategic goals. However, the turn designs of Twitter users who act as the speaker provide an insight into their own expectations of the role that their turns play in subsequent interactions. Contrary to this, their recipients' turn designs precisely indicate how the turns of these Twitter users actually function in the interaction.

Section 5.3 provides a comprehensive analysis of how Twitter users engage in one conflictual interaction extracted from the corpus (see Section 5.3.1) and then reflects on the phenomena that emerge from the analysis (see Section 5.3.2). It finds that although people stereotypically assume that personal attacks are detrimental to the development of interactions, replies that serve as a personal attack play a key role in progressing conflictual Twitter interactions. The investigation of such replies is also of great relevance because it assists in improving the awareness of inharmonious language phenomena among linguists and provides a breakthrough in linguistic research on conflictual interactions. It also provides important insights into the difficulties of identifying linguistic violence in virtual settings, particularly on Twitter.

5.1 Conflictual linguistic strategies in the corpus

This section describes the linguistic strategies used by speakers in conflictual Twitter interactions based on a qualitative analysis of the Conflictual Twitter Interaction Corpus. As linguistic strategies are part of different stages of interactions, this section first sets out the makeup of conflictual Twitter interactions and explains the reasons for the absence of the de-escalation stages in such interactions (see Section 5.1.1). It then identifies the strategies that are used in the escalation stage of conflictual Twitter interactions in the corpus (see Section 5.1.2).

5.1.1 Stages of conflictual Twitter interactions

Existing linguistic research varies in its specific interpretation of the structure of offline and online conflictual interactions (e.g., Eisenberg & Garvey, 1981; Vuchinich, 1990; Hutchby, 2001a; Bousfield, 2007, 2008; Nguyen, 2011; Dobs & Blitvich, 2013; Bou-Franch & Blitvich, 2014). Almost all researchers, however, have noticed that there are two phases with opposing trends throughout conflictual interactions in both online and offline settings. That is, (1) an escalation stage where the intensity of conflict increases and (2) a de-escalation stage where the intensity of conflict decreases. However, none of the conflictual Twitter interactions in the corpus reach the stage of conflict de-escalation.

Conflict management/control is likely to change the status of offline conflictual interactions from escalation to de-escalation, which usually indicates the forthcoming termination of an interaction. For example, Vuchinich (1990) argued that conflictual

interactions terminate in the compromise and resolution of conflict. In the corpus, however, almost all conflictual Twitter interactions end abruptly with one party withdrawing from the interaction. It is impossible for participants to leave a conflictual interaction in face-to-face contexts by disappearing without a closing sequence whereas they can do so on Twitter; they can exit the Twitter App or close the Twitter website. Hence simply withdrawing from Twitter appears to be the quickest and most effective way for users to reduce the impact of conflict on their lives. This suggests that the absence of a closing sequence is a characteristic feature of interactions in Twitter contexts. Although the termination of offline conflictual interactions is usually temporary and the conflict tends to re-escalate afterwards (e.g., Nguyen, 2011), this re-escalation usually occurs after an obvious closing sequence, which is quite a contrast to conflictual Twitter interactions in the corpus because none of these interactions reach the stage of conflict management. Even if one participant actively intervenes in a Twitter conflict, this control is unlikely to have any positive effect: conflictual Twitter interactions as a whole do not de-escalate as a result of this intervention. In terms of social function, researchers have also noticed that the closing sequence in offline conflictual interactions often plays an important role in maintaining social relationships. For example, in a study of conflictual interactions between students in online classes, Xie et al. (2013) found that after the termination of a conflict, students begin to establish new relationships by evaluating both the positive and negative outcomes of the conflict. They also found that the instructor's intervention in students' conflict successfully restored a cycle of social harmony between students. This suggests that the absence of a closing sequence on Twitter may also be due to users' unique perspectives on face-saving, which will be discussed further in Chapter 7 (see Section 7.3).

Twitter interactions, as opposed to face-to-face ones, provide users with a unique interactional environment for engaging in conflict easily and conveniently. First, Twitter enables users to interact via mobile devices at scattered times. Engaging in conflictual Twitter interactions does not take a large chunk of participants' time. They do not have to fear that doing conflict interferes with their normal schedules. Second, escalating conflict on Twitter does not appear to threaten users' physical health because they are not in the same physical space, and conflictual interactions on Twitter are hardly likely to escalate into physical violence as is often the case in face-to-face settings. Third, the anonymity of Twitter may also provide an opportunity for users to express opinions free from the social

constraints that govern face-to-face interactions where participants' expression of everyday emotions is constrained by verbal hygiene measures. This phenomenon was demonstrated in Wang et al.'s (2016) study on weekly trends in work stress and emotion in the US. Wang et al. (2016) found that the expression of negative emotions and stress on Twitter shows a week-long cycle, that is, their amount displays a significant dip on Friday. This contrasts with positive emotion tweets whose amount decreases from Tuesday to Thursday and increases from Friday to Sunday. This may suggest that relieving the daily repressed emotions on Twitter may become a more necessary need for some users on Twitter than in other ways. This need may indirectly also be responsible for Twitter users lacking the avoidance of conflict in interactions.

Governments' and Twitter's relatively lax disciplinary measures against online conflict also allow Twitter users to not much worry about or fear to engage in or observe conflict. In order to protect the right to freedom of expression, most countries do not criminalise the participation in or escalation of online conflictual interactions. Even if they do, participants in online conflictual interactions are rarely convicted due to the difficulty of judging them (see Section 1.3.1 in Chapter 1). Users are not subject to substantial penalties (e.g., fines or criminal liabilities) for escalating conflict on Twitter. Twitter's control over conflictual interactions also typically takes place after the termination of interactions due to the time lag between users' reports and Twitter's measurements. The emergence and development of conflictual interactions are not restricted by Twitter. Due to the protection of the freedom of expression, Twitter does not take any more rigorous measures than deleting posts or suspending accounts ("Twitter—Our range of enforcement options", 2022). Hence participants do not have to bear serious consequences such as threatening their reputations, social status, wealth, physical health in real life, for conducting conflictual interactions on Twitter. For bystanders, they only browse posts in conflictual interactions and solely observe participants' discussions displayed on the screen, without ever tweeting or responding to the posts, and so are even less likely to be punished for the outbreak and/or intensification of conflict. The next section turns the focus to the escalation stage and discusses the conflictual linguistic strategies found in the corpus for escalating conflict in detail.

5.1.2 Escalation strategies

The strategies employed by users during the escalation phase of conflictual Twitter interactions in the corpus were labelled *disagreement*, *personal attack* or *other*, depending on the linguistic behaviours they performed in the interaction as outlined in Section 4.2.4 in Chapter 4. The annotation rules were created based on the review of conflict studies (see Section 2.1 in Chapter 2). *Disagreement* refers to the escalation strategy by which speakers oppose their opponents by objecting to the opponents' viewpoints. *Personal attacks* refer to the escalation strategy by which speakers attack their opponents. Escalation strategies in the corpus that do not fit neatly into either of the two strategies were marked as *Other*, which represents the escalation strategies used by speakers for purposes other than objection or attacks. These annotation rules were tested by conducting a simple survey among Twitter users (see Section 4.2.4 in Chapter 4). This revealed that there are only two escalation strategies in the Conflictual Twitter Interaction Corpus, namely disagreements and personal attacks. No turns were used for purposes other than disagreeing with opponents' views or attacking them.

Strategy 1: disagreements

Disagreements are present in nearly all conflictual Twitter interactions in the corpus, thus becoming the most frequently occurring and widely distributed strategy in the corpus. The analysis shows that there is no obvious difference between disagreements in Twitter and face-to-face contexts. As Stapler stated (1995), when an interaction is characterised by divergent views on a particular matter, speakers defend their views by consciously refuting their opponents' viewpoints. By bringing in additional evidence, disagreements enable speakers to oppose their opponents' viewpoints in a reasonable and logical manner. The analysis shows that Twitter users prefer to express their opinions in conflictual interactions by employing disagreement strategies.

Speakers generally object to their opponents' points of view by partially or totally refuting their arguments. For example, in Excerpt 13, User A5 (Line 08) completely overturns @Brian's opinion on the cause of the Australian bushfires. @Brian (Lines 01–07) updates people on the dire situation of the bushfires in Australia on December 20th, 2019, by posting a photo of the road between Canberra and the Coast. He criticises the views that dissociate the bushfires from climate change but User A5 is convinced that the bushfires are man-made

rather than natural. In order to oppose @Brian, User A5 openly proposes “Arson” (Line 08) as the cause of the bushfires, thus defending their viewpoint.

Excerpt 13 (Interaction 5)

Line	User Account	Text	Level	Turn
01	@Brian	[1R]Australia. This is the road between Canberra and the Coast I think yesterday. The idea that it is somehow not right ‘to talk about #ClimateChange at this difficult time’ is frankly convenient bulllshit. Click on image to see what life is like now for the #firefighters[1WS]. [2R][PIC: https://twitter.com/TwoPaddocks/status/1208207260477882369/photo/1]	tweet	1
02				
03				
04				
05				
06				
07				
08	User A5	[1R]Arson	1st	2

Figure 10

The picture in User A5’s 1st level reply

There are cases where speakers assert that their opponents’ opinions are incorrect by stating that the opponents hold undesirable characteristics, such as User A6’s opposition (Lines 05–07) to @Mary in Excerpt 14. @Mary (Lines 01–04) tweets overtly against the Muslim prayer in the U.S. via a self-made video. In the video, @Mary makes an exaggerated facial expression, mouths the sentence “what the fuck” and covers her face with a black coat, when hearing the sound of Adhan. Although User A6 shares @Mary’s opposition to the Muslim prayer, they question the purpose of her complaint. They argue that it is Muslims’ religious practices, rather than the fact that Muslims disrupt peace, that leads to @Mary’s complaint, “they don’t care if muslims disturb the peace” (Lines 06–07). In this way, User A6 criticises atheists for indiscriminately boycotting religion. User A6 also attributes this boycott to the disappearance of church bells, “Church Bells are a thing of the past due to atheist complaints” (Lines 05–06). It appears that User B6’s attack on @Mary’s identity as an atheist

is sufficient to support their challenge against the purpose of @Mary's complaint. That is, recipients are able to infer from this attack that User A6 opposes @Mary's view.

Excerpt 14 (Interaction 6)

Line	User Account	Text	Level	Turn
01	@Mary	[1R]The intolerable & intrusive sound of #Islamic imperialism in Brooklyn, New York. [2R:ER] [3R]I thought the USA was a #Secular nation... [4R][VIDEO: https://twitter.com/i/status/908092802487132160]	tweet	1
02				
03				
04				
05	User A6	[1R]Church Bells are a thing of the past due to atheist complaints. You can bet they don't care if muslims disturb the peace.	1st	2
06				
07				

Strategy 2: personal attacks

Personal attacks refer to the linguistic strategy that speakers employ to oppose their opponents by, for example, denying, satirising, ridiculing, belittling or mocking them. By using this strategy, speakers intend to attack their opponents, rather than to debate opinion(s) with them. There are 32% (3593) of the replies in the corpus containing personal attacks and Excerpt 15 is one such example.

In Excerpt 15, User B7 (Line 08) attacks User A7 by criticising User A7 for being a furry. The interaction begins with @James's original tweet (Lines 01–04), in which he argues against the rich eating gold chicken wings by demonstrating that it is a waste of both food and money. User A7 (Lines 05–07) questions @James's critique of the rich, arguing that @James is into doctrines, "So you want everyone to starve to death?" (Line 05). They then regard @James as a communist and criticise @James's point of being unrealistic: "communism is only good on paper" (Lines 06–07). Instead of evaluating User A7's critique, User B7 (Lines 08–09) intentionally provokes User A7 by mocking their identity. User B7 implies that User A7 is a furry, "a furry want to talk" (Line 08). User B7's Turn 3 thus deviates from a normal argument of opinion about the golden chicken wing to an attack on User A7's identity. There is no conclusive causal link between User B7's attack on @James's identity and their denial of the reliability of @James's viewpoint.

Excerpt 15 (Interaction 7)

Line	User Account	Text	Level	Turn
01	@James	[1R]I can't even afford food and rich people are eating gold on theirs. Fuck capitalism and fuck the rich [2R][QUOTE: https://twitter.com/InsiderFood/status/1145157506869710848]	tweet	1
02				
03				
04				

(To be continued)

Line	User Account	Text	Level	Turn
05 06 07	User A7	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
08 09	User B7	[1R]Oh so a furry want to talk[Loudly crying face][Loudly crying face][Loudly crying face][Loudly crying face]	2nd	3

In Excerpt 16, User D8 (Lines 28–29) accuses User A8 (i.e., the previous speaker) of being provocative. They do so by comparing their behaviour with that of User A8. User D8 states their support for User A8's opponents, "Someone: *agrees with someone" (Line 28), and criticises User A8's infringement of the opponent's freedom, "Tyrone: 'sToP gEtTiNg InVolVeD'" (Lines 28–29). User A8 (Lines 30–31) defends themself by attacking User D8. They accuse User D8 of having neither the ability to debate, "You sound weak bro" (Line 30), nor the qualification to join this conversation, "You don't need to get involved" (Line 30) and "you're getting in the way" (Line 31). This is followed by User A8's criticism of User D8 for lacking courtesy, "I'm sure your Mum taught you manners, use them" (Lines 31–32).

Excerpt 16 (Interaction 8)

Line	User Account	Text	Level	Turn
28 29	User D8	[1R]Someone: *agrees with someone [2R]Tyrone: "sToP gEtTiNg InVolVeD"	9th	9
30 31 32	User A8	[1R]You sound weak bro. You don't need to get involved, it's a two way convo, and you're getting in the way. I'm sure your Mum taught you manners, use them.	10th	10

The analysis also shows that speakers sometimes construct their turns by using either a disagreement or a personal attack. For example, User A5 objects to @Brian's opinion on the cause of the fire in Excerpt 13. User A6 questions @Mary's ulterior motive for complaining about Muslim prayers in Excerpt 14. User B7 attacks User A7's identity as a furry in Excerpt 15. However, at times, speakers design their turns by combining disagreements and personal attacks. For example, in Excerpt 17, User A1 (Lines 05–07) questions the previous speaker's criticism of the rich eating golden chicken wings for being unrealistic. In response, User B1 attacks User A1's identity, arguing that User A1 is a teenager, an anti-ACAB and a furry, "you're 15, anti-ACAB, a centrist, and a furry" (Line 08). User B1 goes on to openly challenge User A1's opinion for being "completely and utterly invalid" (Line 09).

Excerpt 17 (Interaction 1)

Line	User Account	Text	Level	Turn
05	User A1	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
06				
07				
08	User B1	[1R]you're 15, anti-ACAB, a centrist, and a furry. your opinion is completely and utterly invalid	2nd	3
09				

In cases where disagreements and personal attacks are combined, speakers sometimes place these two strategies in the same reply, as in User B1's reply (Lines 08–09) in Excerpt 17. At times, however, disagreements and personal attacks may be distributed across multiple replies, such as User D3's replies (Lines 29–30 & 31–33) in Excerpt 18. Excerpt 18 displays a fragment of a discussion between Users C3 and D3 about the cause of the Australian bushfires. User C3 opposes User D3 by insisting that there is no link between climate change and drought. Several turns of mutual attacks take place between Users C3 and D3 prior to Turn 10. After being called a loser, User C3 accuses User D3 of defaming them. User D3 hereto attacks User C3 for being thin-skinned, which is challenged by User C3 in Turn 10, "Why thin skinned" (Line 28). User D3 responds to User C3 in two subsequent replies. User D3 starts with an attack on User C3. They provoke User C3 by accusing them of being that kind of person, "Because you think being called a loser is abuse" (Line 29). The provocation is then reinforced by an emoji "😂" (Lines 29–30). In the second reply (Lines 31–33), User D3 questions an example previously provided by User C3, arguing that the example cannot prove that the bushfires are not caused by climate change.

Excerpt 18 (Interaction 3)

Line	User Account	Text	Level	Turn
28	User C3	[1R]Why thin skinned?	9th	10
29	User D3	[1R]Because you think being called a loser is abuse[Face with tears of joy]	10th	11_1
30				
31	User D3	[1R]It doesn't change the fact that your one example is a misquote that got fixed to mean the opposite of what you wanted it to mean	11th	11_2
32				
33				

The findings in this section indicate that Twitter users escalate conflictual Twitter interactions by proposing different opinions to their opponents or attacking them. As Table 4 shows, conflict management aimed at easing the relations between two parties is hardly taken into account by users. One possible explanation might be that they are well aware of the lack of serious penalties for intensifying conflict on the Internet such as ruining reputations, paying fines and bearing criminal responsibilities. Taken together, the results further suggest that Twitter users in conflictual interactions are purportedly more inclined

towards self-expression than to mutual communication. It is unlikely that the primary purpose of their debates is to compensate for the differences in perceptions between themselves and their opponents regarding a certain matter. Engaging in conflict is likely to become an acceptable and effective way for people to present their views unilaterally.

Table 4

A comparison of stages and linguistic strategies in face-to-face interactions and in Twitter interactions

			Offline settings	On Twitter
De-escalation stage	Have or not?		Yes	Not found yet;
	Function		To manage/control or resolve conflict.	\
Escalation stage	Disagreements	Use or not?	Yes, very common.	Yes, very common.
		Function	To object to the opponent's viewpoint.	To object to the opponent's viewpoint.
	Personal attacks	Use or not?	Not sure (lacking linguistic studies, except for those on hate speech).	Yes, common.
		Function	\	To attack the opponent.
	Others	Use or not?	Not sure (lacking linguistic studies).	Not found.
		Function	\	\

5.2 Conflictual linguistic strategies in use

The previous section identified disagreements and personal attacks as the two escalation strategies employed by speakers in the corpus. Although the stereotype argues that divergent opinions do not foster cooperation, the corpus suggests that there is close cooperation between interlocutors because the recipient only issues their turn after a consideration of the (previous) speaker's turn. The recipient, however, does not always seem to understand or interpret the speaker's use of strategies and/or communicative intent in the same way as the speaker, which has also been noticed by Bousfield (2008) when studying linguistic (im)politeness (see Section 3.4.3 in Chapter 3). To address this issue, this section focuses on the interactional functions of conflictual linguistic strategies. Section 5.2.1 specifically examines the ways in which the speaker and the recipient cooperate in conflictual Twitter interactions. Section 5.2.2 then investigates how the recipient in the corpus interprets the (previous) speaker's turn design (including the use of strategies and the purposes of these strategies) and how they develop their own strategic goals. Section 5.2.3 finally compares Twitter users' turn designs with those of their recipients from the perspective of interactional function.

5.2.1 The cooperation between interlocutors

According to DA, to engage in a face-to-face conversation, interlocutors should take turns being the next speaker and issue the next turn. The analysis shows this is also the case in conflictual Twitter interactions because to respond to a post, four steps take place: (1) Twitter users interpret the post about its use of linguistic strategies and its communicative intent; (2) they then decide whether or not to respond to the post and what goals they intend to achieve; (3) they next design their own turns based on their interpretations and strategic goals and (4) they finally issue their own turns. For example, in Excerpt 19 (which is shown as Excerpt 15 in Section 5.1.2), @James's tweet (Lines 01–04) openly criticises the rich eating gold chicken wings as a waste of food and money. User A7 believes that this tweet intends to propagate communism which contradicts their own belief. In order to refuse @James's position and to defend their own, User A7 (Lines 05–07) responds to @James at the 2nd level reply, criticising @James's viewpoint as unrealistic and for being reflective of a communist. After taking a similar approach to User A7, User B7 (Lines 08–09) responds to User A7's criticism. Since User B7 considers User A7's objection to be unsubstantiated and blatantly provocative, they attack User A7, mocking User A7's furry identity.

Excerpt 19 (Interaction 7)

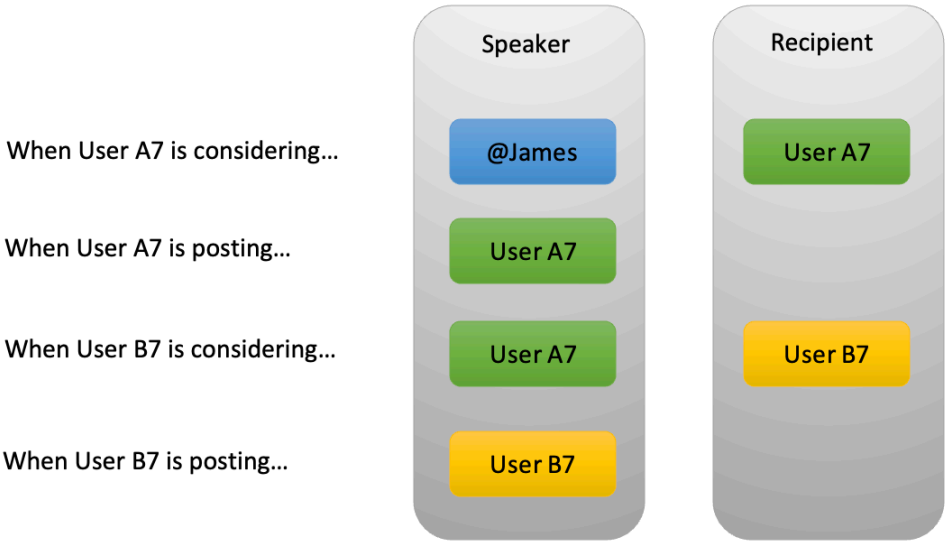
Line	User Account	Text	Level	Turn
01	@James	[1R]I can't even afford food and rich people are eating gold on theirs. Fuck capitalism and fuck the rich [2R][QUOTE: https://twitter.com/InsiderFood/status/1145157506869710848]	tweet	1
02				
03				
04				
05	User A7	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
06				
07				
08	User B7	[1R]Oh so a furry want to talk[Loudly crying face][Loudly crying face][Loudly crying face][Loudly crying face][Loudly crying face]	2nd	3
09				

Twitter users in the corpus appear to be cooperative in conflictual interactions, in contrast to what is stereotypically assumed because they post their own turns successively. For example, in Excerpt 19, User A7 posts Turn 2 based on their interpretation of @James's use of strategies and communicative intent in Turn 1 and also their own strategic goals. Turn 2 is, subsequently, interpreted by User B7 and triggers User B7's Turn 3. As illustrated in Figure 11, the roles of @James, User A7 and User B7 constantly change during this process. Specifically, when User A7 is interpreting @James's turn, @James takes the role of the speaker, and User A7 becomes @James's recipient. Once User A7 posts their turn, their role

shifts to that of the speaker, and accordingly @James assumes the role of the previous speaker. When User B7 is interpreting User A7's Turn 2, User A7 continues to act as the speaker, and User B7 becomes User A7's recipient. Finally, when User B7 posts Turn 3, their role turns to that of the speaker, and User A7's role becomes the previous speaker. The corpus shows that it is through this role shift that Twitter users cooperate with each other in conflictual interactions. It is important to note that in the corpus, this cooperation is not always organised as shown in Excerpt 19 because the user who responds to the speaker may be the same person for whom the speaker's turn is intended.

Figure 11

The role shift in Excerpt 19



Researchers in current interactional studies generally treat APs as the basic unit of analysis (e.g., Brown & Levinson, 1987/1978; Angouri & Locher, 2012; Ben-Menachem & Livnat, 2018; see Section 2.1 in Chapter 2, Sections 3.3 & 3.4 in Chapter 3). They are concerned with how the FFP influences the emergence and construction of the SPP. For example, how does the FPP motivate its recipients to assume the role of the next speaker and to post the SPP? How does the use of linguistic strategies in the FPP affect the next speaker's choice of strategies and the manner in which they express these strategies? For example, in Excerpt 19, since the opinion of @James is opposed to that of User A7, User A7 believes that @James damages their positive face. In order to protect their face, User A7 posts Turn 2 to respond to @James, challenging @James's viewpoint and defending their own position. In the meantime, as @James does not indicate any humiliation of opponents,

User A7 simply treats @James's tweet as an opinion debate rather than a personal attack and constructs their Turn 2 by solely employing a disagreement. This analytical framework based on APs, however, cannot account for the function that the SPP serves in subsequent interactions, such as whether it triggers a response and how it affects the design and realisation of this response. This leads to the need for a third turn. For example, in Excerpt 19, since User B7 considers User A7 to be provocative in Turn 2, they blatantly attack User A7 by mocking their furry identity as a response. This suggests that although User A7 does not employ any personal attack in Turn 2, Turn 2 actually serves as a personal attack in the interaction. Compared to APs, this three-turn system cannot only explain why speakers issue new turns and how they construct them but it is also suitable for examining how speakers' turns actually function in subsequent interactions. This study, therefore, argues that the analysis of the 3rd turn is quite important to reveal the developmental mechanisms of interactions.

While Turn 2 in Excerpt 19 is designed as an objection by User A7, User B7's actions suggest it actually functions as a personal attack. Such inconsistency is pretty common in the corpus: speakers' turns do not serve the roles that speakers expect their turns to serve in the interaction. The next section, therefore, examines the relationship between the speaker's use of strategies and the recipient's actions in response to these strategies in more detail.

5.2.2 The speaker's use of strategies vs. the recipient's responses

Two situations emerged from the analysis. First, the linguistic behaviours performed by the recipient are consistent with the (previous) speaker's use of strategies. Second, mismatches occur between the speaker's strategic choices and the recipient's responses. These are discussed in detail in turn below.

Situation 1: the recipient's actions correspond to the (previous) speaker's use of strategies

In conflictual interactions in which the speaker designs their turn to consist only of one type of strategy, either disagreements or personal attacks, the recipient sometimes responds to the speaker with the same strategy. This fact usually suggests that the recipient has not only fully understood the strategies used by the (previous) speaker and successfully predicted the speaker's communicative intent but has also decided to go along with their interpretation. For example, in Excerpt 20, in order to provoke the public into resisting vegan

lifestyles, @Susan (Lines 01–03) indicates that vegan food is unpalatable. She illustrates it with a picture, showing that all non-vegan food was sold out during the hurricane while the shelves with vegan food remained almost fully stocked, “Even in a disaster no one wants the vegan food” (Line 01). User A9 interprets @Susan’s tweet as an attack on vegans and responds to it in accordance with this interpretation. They defend vegans by arguing that vegan lifestyles are the most effective method to prevent natural disasters, “if everyone were eating this food, the natural disaster may not have even happened” (Lines 04–05). Although this sentence appears to serve as a reasonable rebuttal to @Susan’s argument, in doing so, User A9 attacks @Susan off-record. They imply that people like @Susan who resist vegan food are the problem, which suggests that @Susan’s tweet is perceived by User A9 as an attack.

Excerpt 20 (Interaction 9)

Line	User Account	Text	Level	Turn
01 02 03	@Susan	[1R]Even in a disaster no one wants the vegan food. [2R][PIC: https://twitter.com/hillisthekillis/status/903720326231719936/photo/1]	tweet	1
04 05 06	User A9	[1R]It’s ironic because if everyone were eating this food, the natural distaster may not have even happened[Woman shrugging]	1st	2

Figure 12

The picture in @Susan’s tweet

Another case in point is Excerpt 21. In order to oppose User A7 (Lines 05–07), User D7 (Lines 08–09) implies that User A7 is not qualified to express views. They do this by identifying User A7 as a furry, “Oh so a furry want to talk” (Line 08). User A7 reacts to the blatant provocation. They take issue with D7’s use of furries, “Please don’t pin this on furries” (Line 10), and defends themself. They argue that User D7’s attack stems from their

prejudice against the group of furries, “Everyone is susceptible to propaganda and it’s completely unrelated to the fandom” (Lines 10–11). As a result, User D7 is criticised for being narrow-minded.

Excerpt 21 (Interaction 7)

Line	User Account	Text	Level	Turn
05 06 07	User A7	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
08 09	User D7	[1R]Oh so a furry want to talk[Loudly crying face][Loudly crying face][Loudly crying face][Loudly crying face]	2nd	9
10 11	User A7	[1R]Please don’t pin this on furries. Everyone is susceptible to propaganda and it’s completely unrelated to the fandom.	3rd	10

In Excerpt 22, User A8 (Lines 30–32) attacks User D8, arguing that User D8 is not competent or qualified to ‘talk’ in the conversation, “You sound weak bro. You don’t need to get involved” (Line 30) and “you’re getting in the way” (Line 31). User A8 goes on to put User D8 down by treating them like a naughty child, *“I’m sure your Mum taught you manners, use them”* (Lines 31–32). User D8 interprets User A8’s Turn 10 as an attack and attacks User A8 in response. They first humiliate User A8’s mother, “I’m sure your mum didn’t teach you manners” (Line 33). This is followed by their critique of User A8 for being intolerant and unsympathetic, “you’d rather people die from a virus than swallow your pride” (Lines 34–35). User D8 finalises the reply by implying that it is User A8 who is not entitled to express opinions, “Twitter is a free platform” (Line 35). This sentence also shows their strong sarcasm towards User A8 due to the insertion of an emoji of a thumbs up, “👍” (Line 35).

Excerpt 22 (Interaction 8)

Line	User Account	Text	Level	Turn
28 29	User D8	[1R]Someone: *agrees with someone [2R]Tyrone: “sToP gEtTiNg InVolVeD”	9th	9
30 31 32	User A8	[1R]You sound weak bro. You don’t need to get involved, it’s a two way convo, and you’re getting in the way. I’m sure your Mum taught you manners, use them.	10th	10
33 34 35	User D8	[1R]I’m sure your mum didn’t teach you manners because you’d rather people die from a virus than swallow your pride... [2R]Twitter is a free platform[Thumbs up]	11th	11

Situation 2: the recipient’s responses contradict the (previous) speaker’s use of strategies

The corpus also suggests that mismatches occasionally occur between the speaker’s use of strategies and the recipient’s responses, in both cases where the speaker uses one strategy, i.e., a disagreement or a personal attack, or combines two types of strategies. This

study argues that there may be two possible reasons for such mismatches: first, the recipient does not adequately identify the (previous) speaker's strategies or speculate on the speaker's strategic goals. Second, although the recipient fully interprets the (previous) speaker's strategic design, they do not construct their own turn along with their interpretation. Because the further identification of these two reasons requires a judgement of the recipient's innermost ideas, which is beyond the focus of this study on communicative intent, the study does not distinguish the specific reasons for those mismatches. Both possibilities, however, result in the recipient failing to respond to the (previous) speaker in the manner that the speaker expects. For example, in Excerpt 23, @Jessica (Lines 01–04) posts a question to women, "What would you do if all men had a 9pm curfew?" (Lines 01–02), and makes her opinion on gender equality public through adding hashtags (e.g., #feminism #maleprivilege). In response, User A10, however, openly denies the legitimacy of @Jessica's insinuation, "I would protest it – because 99% of men aren't the animals you are trying to convince everyone they are" (Lines 05–06). They then accuse @Jessica of being "sexist" (Line 07). This suggests that from the perspective of User A10, @Jessica is deliberately inciting gender hatred.

Excerpt 23 (Interaction 10)

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
05	User A10	[1R]I would protest it – because 99% of men aren't the animals you are trying to convince everyone they are. Stop being sexist. #twowaystreet	1st	2
06				
07				

In Excerpt 24, @PinkNews (Lines 01–02) challenges the stereotype about make-up and gender, by citing a video of a young boy demonstrating his superior make-up skills. This is supported by User A11 (Line 03). As the boy is underage, the posts of @PinkNews and User A11 are considered to be criminal by User B11. User B11 (Lines 04–05) verbalises this view by accusing them of being paedophiles, "if you think pedophilia and child abuse is cute" (Line 04). User B11's threats against them are subsequently followed, "In which case, I'd call the FBI on ya" (Line 05). As can be seen, although @PinkNews and User A11 clarify their posts as promoting gender equality, their posts are still perceived as an attack.

Excerpt 24 (Interaction 11)

Line	User Account	Text	Level	Turn
01 02	@PinkNews	[1R]Makeup has no gender[Person tipping hand] [2R][VIDEO: https://twitter.com/i/status/1287036885953650688]	tweet	1
03	User A11	[1R]awww so cute!	1st	2
04 05	User B11	[1R]Yeah, if you think pedophilia and child abuse is cute [2R:ER] [3R]In which case, I'd call the FBI on ya	2nd	3

Another case in point is Extract 25 that shows a discussion about whether Logan Paul, who previously taunted the dead in the Aokigahara forest, can really help the Amazon fires as he states in the tweet. After criticising the previous speaker for being unforgiving, "Why are you assuming he would… Visit the forest" (Line 23), User B12 turns to object to the previous speaker's view. They first argue that everyone is responsible for protecting the Amazon rainforest, "we all need to do our part, no matter how shitty of a person we are" (Lines 24–25). This argument is followed with an example of how Logan Paul is able to raise funds for the fire through his influence, "Famous people SHOULD be putting their money/celebrity towards fighting this. His fan base will follow his lead" (Lines 25–27). In the second 5th level reply, User B12 attacks the previous speaker once again. They challenge the previous speaker's motive for accusing Logan Paul, "because you dislike LP you want to hinder helping the planet" (Lines 52–53). They then further provoke the previous speaker, asserting that the previous speaker intends to exterminate humanity, "Can't wait for global warming to kill us all" (Line 53). User C12 only responds to User B10's first reply. They neither imply that they are under attack by defending themself nor do they openly counter User B12. User C12 only opposes User B12 by insisting that Logan Paul is definitely dangerous to the Amazon rainforest, "it's a youtuber known for taking things to the extreme" (Lines 28–29). This suggests that User C12 does not take the attack from User B12 seriously.

Excerpt 25 (Interaction 12)

Line	User Account	Text	Level	Turn
23 24 25 26 27	User B12	[1R]Why are you assuming he would... Visit the forest. I'm not going to argue with you, we all need to do our part, no matter how shitty of a person we are. Famous people SHOULD be putting their money/celebrity towards fighting this. His fan base will follow his lead	4th	5_1
28 29	User C12	[1R]Because it's a youtuber known for taking things to the extreme, just like his brother (though in different ways).	5th	6
Omit Lines 30–49				

(To be continued)

Line	User Account	Text	Level	Turn
50	User B12	[1R]What more can he do? He is literally asking what he can do in the tweet you responded to. If this were anyone else you would even notice but because you dislike LP you want to hinder helping the planet. Can't wait for global warming to kill us all	5th	5_2
51				
52				
53				

5.2.3 The function of turn design

According to the analysis in the previous sections, Twitter users' turns do not merely serve as an adaptation of the previous speaker's turn as DA suggests. More importantly, they are the outcome of cooperation among Twitter users and reflect their strategic designs of their own turns, in response to the previous speaker's contribution. In conflictual interactions, Twitter users constantly shift from being a recipient to being a speaker. They first take the role of the recipient of the (previous) speaker's turn and interpret the design of this turn. They need to identify what types of strategies are employed by the speaker to speculate about what functions the speaker intends their turn to serve. They also need to determine whether or not to withdraw from the interaction and whether or not to respond to the speaker alongside their interpretation. After interpreting the (previous) speaker's use of strategies and communicative intent, Twitter users communicate their assessment of the speaker's turn design and their own strategic goals within the interaction by publishing their own responses in the next turn and allow their posts to be freely interpreted by their recipients. Their role thus shifts from the recipient to that of the (next) speaker.

However, as Section 5.2.2 found, Twitter users' turns may not contribute to the strategic purposes for which they expect them to do in interactions. Thus, Twitter users' turn designs and those of their recipients provide equally important practical implications for revealing the developmental mechanisms of conflictual interactions. The constructions of Twitter users' turns reflect the role that their turns are expected to play in ensuing conflictual interactions by themselves. Their recipients' turn constructions are precisely the function that these Twitter users' turns actually serve as in the interaction.

For example, in Excerpt 23, @Jessica aims to call for gender equality while avoiding being perceived as sexist. They do this by emphasising that the audience consists of women, "Ladies" (Line 01), and by reminding men to read carefully before commenting, "Dudes: Read the replies and pay attention" (Lines 02–03). User A10 completely ignores @Jessica's purpose and rather regards @Jessica's turn as a personal attack. They accuse @Jessica of deliberately inciting gender hatred and request @Jessica to "Stop being sexist" (Lines 06–

07). User A10's response suggests that @Jessica's turn actually functions as an attack rather than an objection in the interaction. In Excerpt 24, instead of attacking any individual or group, @PinkNews only challenges the stereotype about makeup and gender, "Makeup has no gender" (Line 01), so as to call for gender equality. However, User B11 considers @PinkNews's promotion of gender equality as a crime against the underage boy and overtly attacks @PinkNews. They accuse @PinkNews of being a paedophile, which suggests that @PinkNews' turn actually serves as a personal attack in the interaction. Another case in point is in Extract 25. In order to call for a proper treatment of the celebrity effect, User B12 suggests that "Famous people SHOULD be putting their money/celebrity towards fighting this" (Lines 25–26), by explaining that "His fan base will follow his lead" (Lines 26–27). In the meantime, in order to deal with the previous speaker's attack, User B12 also accuses the previous speaker of being unforgiving, "Why are you assuming he would… Visit the forest" (Line 23), and provokes them, "Can't wait for global warming to kill us all" (Line 53). However, since User C12 does not deal with User B12's attack and instead explains why Logan Paul cannot serve as a positive role model, User B12's turn actually acts as a disagreement in the interaction. This suggests that this study should not only investigate how Twitter users construct their turns by using conflictual linguistic strategies, but it should also focus on the interactional functions of their turns by examining the turn designs of their recipients' turns.

Together, the results in this section provide an important insight into turn design in conflictual Twitter interactions at a theoretical level. They demonstrate that the conversational mechanism maintained by DA still applies to conflictual Twitter interactions because users need to constantly switch their roles in interactions to issue the next turn, thus cooperating with each other. They first, as the recipient, interpret the (previous) speaker's turn design, then publish the next turn as the (next) speaker based on their interpretations and strategic goals and finally leave their turns open for interpretations by their recipients. However, there may be a mismatch between the designs of their own turns and those of their recipients' turns. Since these two types indicate the strategic goals and practical roles of Twitter users' turns, respectively, they are equally important to reveal the developmental mechanisms of conflictual Twitter interactions.

5.3 An example of using strategies

Section 5.1 discussed how Twitter users construct their turns by using disagreements and personal attacks. Section 5.2 explained how Twitter users select linguistic strategies and formulate their strategic goals, examined how these strategies and goals are interpreted by their recipients and also considered both the practical implications of Twitter users' and their recipients' turn designs. Based on the above findings, this section presents a comprehensive analysis and some reflections on how Twitter users engage in conflictual Twitter interactions. Specifically, Section 5.3.1 analyses an interaction from the corpus (i.e., Interaction 2) in detail. It focuses on how speakers construct their turns by interpreting the strategies used in the previous turn and the purposes of these strategies and how their turns drive the presentation of the next turn. Section 5.3.2 discusses the phenomena and issues evident in the analysis, particularly with regard to the relationship between language strategies and conflictual interactions.

5.3.1 Case study

Interaction 2 originates from a tweet posted by @John on September 1st, 2019. @John tweets a video with a comment, "A desperate cry for help 😢 😢 😢 😢 #HurricaneDorian #Abaco #Bahamas Lord please help us" (Lines 01–03). In the video, a Bahamian woman cries out about the desperate situation that the Bahamian people are facing following Hurricane Dorian and asks people to pray for them, by filming how Hurricane Dorian is severely damaging their homes. This section only focuses on a part of Interaction 2 that is shown in Excerpt 26, mainly the interaction between Users B2 and C2. As Figure 13 shows, it ranges from the original tweet to the 10th level reply, including 21 turns and involving five participants: @John, User A2, User B2, User C2, User D2. @John, User A2 and User B2 are theists; Users D2 and C2 are atheists.

Excerpt 26 (Interaction 2)

Line	User Account	Text	Level	Turn
01 02 03 04	@John	1R]A desperate cry for help[Crying face][Crying face][Crying face][Crying face]#HurricaneDorian #Abaco #Bahamas Lord please help us [2R][VIDEO: https://twitter.com/i/status/1168227588533293056]	tweet	1

(To be continued)

Line	User Account	Text	Level	Turn
05 06 07 08 09	User A2	[1R]Father in Heaven, please safeguard Your people in the path of Hurricane Dorian. Loose Your mighty angels to minister unto each of them at this time. Fill them with Your peace that passes all understanding. Keep them together in love. In Jesus' Name, Amen.	1st	2
10 11	User B2	[1R]And the animals and wildlife too, Lord, created by you.[Folded hands][Folded hands][Folded hands]	2nd	3
12 13 14 15	User C2	[1R]I'm assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn't sound like a very compassionate deity to me.	3rd	4
16 17 18 19	User B2	[1R]Relax Atheist, your day of reckoning will come... everyone's does. But it's funny how, when one's life or a loved one's life is in danger people turn to asking their God to save them. But that would never be you, right??? You'll have to wait and see...	4th	5_1
20 21 22 23	User C2	[1R]You are quite correct - it is funny how people to turn to a god to save them from bad things whilst never wondering why their loving all-powerful god allows the things to happen in the first place. The hypocrisy baffles me.	5th	6_1
Omit Lines 24–30[22]				
31 32	User C2	[1R]Presumably they'll use the all encompassing get out clause that their god moves in mysterious ways.... hmmm....	5th	6_2
33 34 35 36 37	User C2	[1R]Oh, and by the way - obviously I hope that those that have been affected by the storm are ok - because I'm not a horrible bellend.... I just don't see that praying is likely to help them in any way whatsoever. I was brought up with religion and it has done nothing for me.	5th	6_3
38 39 40 41	User B2	[1R]Then why troll the comments of those that do believe? ... have you nothing better to do, that can help people that are afraid and in danger of dying in a hurricane, who are asking us to pray for them?[1WS] This is your solution?	6th	9
42 43 44 45 46	User C2	[1R]I've not trolled anyone. I've not said anything insulting about any individuals. Just making the point that I fail to see how it will actually help anyone. They are better supported by actual action rather than prayers - whether it be financial or structural aid. Just my opinion.	7th	10_1
47 48 49 50	User B2	[1R]Sweetie, these people in the videos where the sorm is now can't get help until it's over. THEY are asking for prayers bc they are scared.[1WS] Do this for them, it's called COMPASSION, and it won't hurt you!!![Folded hands]	8th	11
51 52 53 54 55	User C2	[1R]I can assure you I'm thinking of them, & I'm hoping that they are ok. I just won't pray to a god I don't believe in. I apologise if I caused you offence - I can assure you it wasn't intended that way. I just keep seeing posts about praying, yet bad stuff still happens in droves.	9th	12
56 57 58 59 60	User B2	[1R]Unfortunately the world is made up of good and bad things. We are also given free will to chose a given path bc God didn't make us to be mindless stick figures to experience life in a bubble.[1WS] We make mistakes, and sometimes we need help, or we're thankful.	10th	13_1

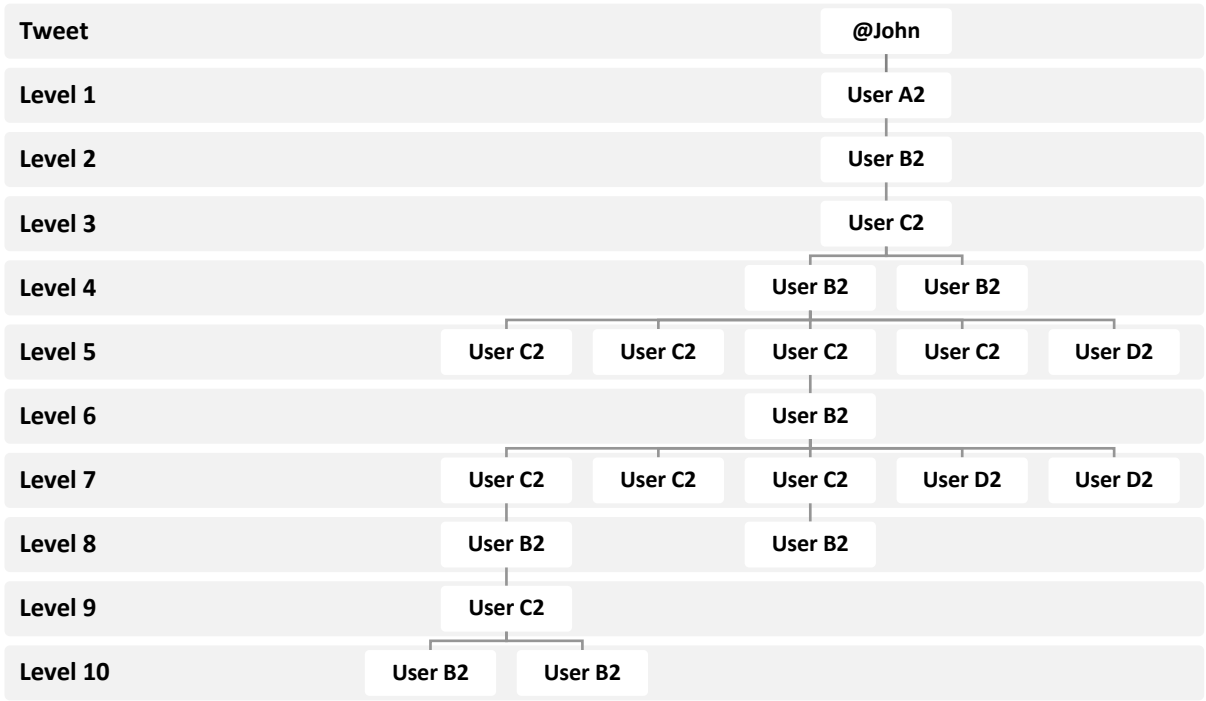
(To be continued)

[22] The omitted part is not the interaction between Users B2 and C2

Line	User Account	Text	Level	Turn
61 62 63 64 65	User B2	[1R]... so we give thanks to him. It's a comforting way to live in faith. Especially as we realise we're not in control of ANYTHING!Someone once said, even if there's a chance there is no God, living in faith is a nice way to live this life we're given-It's true	10th	13_2
66 67 68	User C2	[1R]I'd happily donate - and often do - to support those in peril or danger. That surely benefits them more than me telling them I'm thinking of them???	7th	10_2
Omit Lines 69–70				
71 72	User C2	[1R].... and to be honest - you threatened me with a day of reckoning!	7th	10_3
73 74 75	User B2	[1R]No. I didn't "threaten" you with a day of reckoning... it's just a fact of life, it happens to everyone one day. You will learn what your honest reactions are when it happens to you..	8th	15
76	User D2	[1R]They aren't asking you to pray for them. FFS, piss off.	7th	16_1
77	User D2	[1R]Your prayers help no one.	7th	16_2
78 79 80 81 82	User C2	[1R]Everyone has their right to an opinion - believers and non-believers. If you believe, then that is obviously fine. I just don't understand why people look to something to save them, when if their beliefs are correct, that god should prevent the bad stuff in the 1st place.	5th	6_4
83 84	User D2	[1R]I would never ask for any God's help if I was in danger as god does not exist. Bloody bible-thumping weirdo.	5th	17
Omit Lines 85–96				
97 98 99	User B2	[1R]And whilst you sit there "assuming", what exactly did you do for the girl in the video who's PLEADING FOR PRAYERS???? If you don't want to pray for her... THEN DON'Tperiod[Neutral face]	4th	5_2

Figure 13

The flowchart of Excerpt 26



User A2 (Lines 05–09) at the 1st level prays for God's blessing of the people who are suffering due to Hurricane Dorian, "please safeguard Your people in the path of Hurricane Dorian" (Lines 05–06), as expected by @John and the Bahamian woman in the video. User B2 (Lines 10–11), then, supports @John and User A2 by expanding the scope of prayer to animals and wildlife, "And the animals and wildlife too" (Line 10).

Excerpt 27 (Interaction 2)

Line	User Account	Text	Level	Turn
05	User A2	[1R]Father in Heaven, please safeguard Your people in the path of Hurricane Dorian. Loose Your mighty angels to minister unto each of them at this time. Fill them with Your peace that passes all understanding. Keep them together in love. In Jesus' Name, Amen.	1st	2
06				
07				
08				
09				
10	User B2	[1R]And the animals and wildlife too, Lord, created by you.[Folded hands][Folded hands][Folded hands]	2nd	3
11				

In order to object to User B2's view, User C2 (Lines 12–15) implies the non-existence of God by saying that God's failure to stop Hurricane Dorian suggests God's lack of compassion. User C2 first suggests that God is not omnipotent by pointing out a logical error of theists through a personal speculation, "I'm assuming that" (Line 12). In the meantime, they also cast Christians as stupid and blinded, "no-one is considering" (Line 13). They follow this up with an explicit criticism of God's lack of compassion, "Doesn't sound like a very compassionate deity" (Lines 14–15), but they also redress it by stating that this is only a personal view, "to me" (Line 15). User C2's reply, however, is interpreted by User B2 as a personal attack in Excerpt 29.

Excerpt 28 (Interaction 2)

Line	User Account	Text	Level	Turn
12	User C2	[1R]I'm assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn't sound like a very compassionate deity to me.	3rd	4
13				
14				
15				

In response, User B2 links their two 4th level replies through a coordinating conjunction, "And" (Line 97), to establish a reasonable premise for their objection to User C2's point. User B2, in the first reply, begins with threatening User C2 by asserting that User C2 will one day abandon their atheistic views out of desperation, "Relax Atheist, your day of reckoning will come… everyone's does" (Lines 16–17) and "that would never be you, right???" (Line 19). They then object to User C2's atheistic view by elaborating that the significance of prayer is to provide spiritual support rather than material assistance to

people in desperate situations, “when one’s life or a loved one’s life is in danger people turn to asking their God to save them” (Lines 17–18). In the second reply, User B2 (Lines 97–99) continues to oppose User C2 by demonstrating why User C2’s denial lacks credibility. They question User C2 for being provocative and not truly attempting to help the woman in the video at all. They also criticise User C2 for being narrow minded by asking User C2 not to interfere with others’ freedom of belief to be an atheist. In addition, apart from provoking User C2 through threats, User B2 remains generally restrained in their logically relevant refutation of User C2’s criticism of prayer. Instead of publicly rejecting User C2’s viewpoint, User B2 leads User C2 to recognise their problem by asking “what exactly did you do for the girl in the video” (Lines 97–98). When requesting User C2 to not interfere with others’ freedom of belief, User B2 also expresses their respect for User C2’s identity as an atheist by offering an option of not praying, “If you don’t want to pray for her…” (Lines 98–99). This demonstrates that User B2 attempts to avoid escalating the conflict consciously by protecting User C2’s face. It is also evident through User C2’s response that User B2’s intent in conflict control has been successfully communicated to D2 (see Excerpt 30) and Users C2 (see Excerpt 31).

Excerpt 29 (Interaction 2)

Line	User Account	Text	Level	Turn
16	User B2	[1R]Relax Atheist, your day of reckoning will come... everyone’s does. But it’s funny how, when one’s life or a loved one’s life is in danger people turn to asking their God to save them. But that would never be you, right??? You’ll have to wait and see...	4th	5_1
17				
18				
19				
Omit Lines 20–96				
97	User B2	[1R]And whilst you sit there “assuming”, what exactly did you do for the girl in the video who’s PLEADING FOR PRAYERS???? If you don’t want to pray for her... THEN DON’Tperiod[Neutral face]	4th	5_2
98				
99				

Both Users C2 and D2 counter User B2’s threat to atheists. User D2 opposes User B2 by directly rejecting User B2’s threat, “I would never ask for any God’s help” (Line 83), and publicly stating that “god does not exist” (Lines 83–84). User D2 further undermines the validity of theists by suggesting that the promotion of theistic views is fraught with ulterior motives, “Bloody bible-thumping weirdo” (Line 84). This constitutes User D2’ attack on User B2.

Excerpt 30 (Interaction 2)

Line	User Account	Text	Level	Turn
83	User D2	[1R]I would never ask for any God’s help if I was in danger as god does not exist. Bloody bible-thumping weirdo.	5th	17
84				

At the 5th level, User C2 also immediately counters User B2's challenge with three replies. In the first two replies, they acknowledge that theists always put their hope in God by arguing that "you are correct" (Line 20). In the light of theists' blind faith in God and denial of the fact that God cannot prevent disasters, User C2 then criticises the theists' (including User B2) logical thinking abilities. They finalise the first reply by criticising the theists for being hypocrites, "The hypocrisy baffles me" (Line 23), thus denying the value of prayer. User C2, subsequently, defends themself against User B2's threat, "obviously I hope that those that have been affected by the storm are ok" (Lines 33–34). They also provide a specific reason why prayer is futile by arguing that God cannot provide substantial help for the victims, "don't see that praying is likely to help them in any way whatsoever" (Lines 35–36). This is supported by presenting User C2's own experience, "I was brought up with religion and it has done nothing for me" (Lines 36–37).

Excerpt 31 (Interaction 2)

Line	User Account	Text	Level	Turn
20 21 22 23	User C2	[1R]You are quite correct - it is funny how people to turn to a god to save them from bad things whilst never wondering why their loving all-powerful god allows the things to happen in the first place. The hypocrisy baffles me.	5th	6_1
Omit Lines 24–30				
31 32	User C2	[1R]Presumably they'll use the all encompassing get out clause that their god moves in mysterious ways.... hmmm....	5th	6_2
33 34 35 36 37	User C2	[1R]Oh, and by the way - obviously I hope that those that have been affected by the storm are ok - because I'm not a horrible bellend.... I just don't see that praying is likely to help them in any way whatsoever. I was brought up with religion and it has done nothing for me.	5th	6_3

Although User C2 intends to construct their responses as disagreements after sensing User B2's control over conflict, not every reply of User C2 is interpreted as a disagreement by the recipient. User C2's third 5th level reply is perceived by User B2 as a personal attack, triggers sustained personal attacks and sets the stage for the ensuing escalation of conflict. User B2 counters User C2's personal attack by blatantly accusing User C2 of being a troll, "why troll the comments of those that do believe?" (Line 38), and denouncing User C2 for not really wanting to offer help for the victims. User B2's overt attack on User C2 results in Users D2 and C2 cooperating against User B2, thus escalating the individual conflict between Users B2 and C2 into an antagonism between the atheistic and theistic camps in a real sense (see Excepts 33 & 34).

Excerpt 32 (Interaction 2)

Line	User Account	Text	Level	Turn
38 39 40 41	User B2	[1R]Then why troll the comments of those that do believe? ... have you nothing better to do, that can help people that are afraid and in danger of dying in a hurricane, who are asking us to pray for them?[1WS] This is your solution?	6th	9

User C2 first defends themself against User B2's overt attack by arguing that "I've not trolled anyone. I've not said anything insulting about any individual" (Lines 42–43). User C2, subsequently, reiterates their point about the futility of prayer, "They are better supported by actual action rather than prayers" (Lines 44–45). Following this, a possible solution is provided to deepen their opinion, "I'd happily donate" (Line 66). User C2 ends their turn by explicitly accusing User B2 of being the one who attacks others, "you threatened me" (Line 71). It was found that User C2's reaction is entirely a self-defence triggered by User B2's personal attack, but User C2's third reply is still treated as an attack by User B2 (see Excerpt 36).

Excerpt 33 (Interaction 2)

Line	User Account	Text	Level	Turn
42 43 44 45 46	User C2	[1R]I've not trolled anyone. I've not said anything insulting about any individuals. Just making the point that I fail to see how it will actually help anyone. They are better supported by actual action rather than prayers - whether it be financial or structural aid. Just my opinion.	7th	10_1
Omit Lines 47–67				
66 67 68	User C2	[1R]I'd happily donate - and often do - to support those in peril or danger. That surely benefits them more than me telling them I'm thinking of them???	7th	10_2
Omit Lines 69–70				
71 72	User C2	[1R].... and to be honest - you threatened me with a day of reckoning!	7th	10_3

User D2 also responds to User B2's attack at the 7th level. This is the second time that User D2 has intervened in the two-participant interaction between Users B2 and C2. From the 4th level onwards, Users B2 and C2 explicitly indicate that they are both recipients and commenters of each other's replies by repeatedly using the second personal pronoun, "you" (e.g., Lines 19, 20, 71 & 98). User D2 intervenes in Users B2 and C2's interaction at the 5th level for the first time (see Lines 83–84 in Excerpt 30), which results in a two-participant interaction between Users B2 and D2. Afterwards, User D2, however, leaves their interaction with User B2 and returns to the interaction between Users B2 and C2 again in Excerpt 34. User D2 does this by countering User B2's personal attack together with User C2 at the 7th level. Unlike User C2, User D2 proactively attacks User B2 by challenging User B2's prayer as

both unfounded and unhelpful. User D2 also escalates their attack by upgrading their word choice to swear words. They convey their extreme annoyance not only by using "FFS" (Line 76) but also by denying User B2 the right to express opinions, "piss off" (Line 76). Thus, an antagonism between the theistic and atheistic camps is formally established.

Excerpt 34 (Interaction 2)

Line	User Account	Text	Level	Turn
76	User D2	[1R]They aren't asking you to pray for them. FFS, piss off.	7th	16_1
77	User D2	[1R]Your prayers help no one.	7th	16_2

Before User B2 has a chance to reply to User C2's personal attack at the 7th level, User C2 returns to the previous conversation with User B2. They abandon the personal attacks just used in their 7th level replies (see Excerpt 33) and put forward the fourth 5th level reply (Lines 78–82). User C2 take it as a compliment against User B2's (Lines 16–19) turn at the 4th level (see Excerpt 29) to further clarify their defensive purpose in a less offensive manner. User C2 implies that User B2 interferes with others' freedom of speech and denies the validity of the theists' beliefs by arguing that prayer to God is meaningless. The discussion thus returns to the original atheistic viewpoint about the non-existence of God. First of all, User C2 implies that User B2's view is not superior to others' views by claiming that "Everyone has their right to an opinion – believers and non-believers" (Lines 78–79). They go on to challenge the validity of the theists' beliefs by questioning the value of the prayer, "when if their beliefs are correct, that god should prevent the bad stuff in the 1st place" (Lines 80–82), thus denying the value of the existence of God. In their expression, User C2 also attempts to mitigate the damage to User B2's positive face. They recognise User B2's freedom of belief by arguing that "If you believe, then that is obviously fine" (Line 79) and by personalising their objection, "I just don't understand" (Lines 79–80), which suggests that User C2 attempts to redress the face-threat to User B2 in their previous replies.

Excerpt 35 (Interaction 2)

Line	User Account	Text	Level	Turn
78 79 80 81 82	User C2	[1R]Everyone has their right to an opinion - believers and non-believers. If you believe, then that is obviously fine. I just don't understand why people look to something to save them, when if their beliefs are correct, that god should prevent the bad stuff in the 1st place.	5th	6_4

User B2 does not react to User D2's blatant attack in Excerpt 34 or User C2's objection in Excerpt 35 but only selectively responds to User C2's first 7th level reply and third 7th level reply (see Lines 42–46 & 71–72 in Excerpt 33). In Excerpt 36, User B2 first refuses User C2's self-defence and clarification of opinions by giving their reason for praying, "THEY are asking for prayers" (Line 48). They then openly command User C2 to pray, "Do this for them" (Line 49). In response to User C2's accusation in the third reply, User B2 defends themself, arguing that "I didn't threaten you" (Line 73). They then threaten User C2 again as they did at the 4th level (see Excerpt 29) by arguing that they are only stating a fact.

Excerpt 36 (Interaction 2)

Line	User Account	Text	Level	Turn
47	User B2	[1R]Sweetie, these people in the videos where the sorm is now can't get help until it's over. THEY are asking for prayers bc they are scared.[1WS] Do this for them, it's called COMPASSION, and it won't hurt you!!![Folded hands]	8th	11
48				
49				
50				
Omit Lines 51–72				
73	User B2	[1R]No. I didn't "threaten" you with a day of reckoning... it's just a fact of life, it happens to everyone one day. You will learn what your honest reactions are when it happens to you..	8th	15
74				
75				

Only User B2's Turn 11, which is treated as an attack, receives a response from User C2. User C2 defends themself against User B2's threat by arguing that "I'm thinking of them, & I'm hoping that they are ok" (Lines 51–52) and rejects User B2's assertion again by arguing that "I just won't pray to a God" (Line 52). User C2, then, suddenly shifts to apologise for possible attacks on User B2 by stating that "I apologise if I caused you offence" (Lines 52–53). This allows User B2 to abandon their aggressive personal attacks.

Excerpt 37 (Interaction 2)

Line	User Account	Text	Level	Turn
51	User C2	[1R]I can assure you I'm thinking of them, & I'm hoping that they are ok. I just won't pray to a god I don't believe in. I apologise if I caused you offence - I can assure you it wasn't intended that way. I just keep seeing posts about praying, yet bad stuff still happens in droves.	9th	12
52				
53				
54				
55				

Although User B2 continues to oppose User C2 by insisting on their theistic viewpoint (Lines 56–65), they eventually terminate their interaction by establishing common ground with User C11. User B2 eludes the roots of the divergence between theists and atheists and instead points out that faith is equally important to atheists as to theists. However, the mitigation is temporary, as Users B2 and C2 continue to attack each other after Turn 13. Since their interaction afterwards occurs in another interactive stream triggered by a new

user's reply at the 4th level (i.e., User E2), it is not analysed in detail here. A discussion of this part can be found in Excerpt 127 (see Section 7.3 in Chapter 7).

Excerpt 38 (Interaction 2)

Line	User Account	Text	Level	Turn
56	User B2	[1R]Unfortunately the world is made up of good and bad things. We are also given free will to chose a given path bc God didn't make us to be mindless stick figures to experience life in a bubble.[1WS] We make mistakes, and sometimes we need help, or we're thankful.	10th	13_1
57				
58				
59				
60				
61	User B2	[1R]... so we give thanks to him. It's a comforting way to live in faith. Especially as we realise we're not in control of ANYTHING!Someone once said, even if there's a chance there is no God, living in faith is a nice way to live this life we're given- It's true	10th	13_2
62				
63				
64				
65				

5.3.2 Personal attacks in the case

The common perception of personal attacks is that they impede the progress of interactions because they prevent logical objections from being raised. This stereotype emerges since using personal attacks does not contribute to speakers' reasonable rebuttals of their opponents' viewpoints. Personal attacks are, thus, often seen as being detrimental to the further development of an argument between interlocutors as they easily turn a conflictual interaction into a name-calling war or lead to an abrupt end. However, based on the discussion in Section 5.2, there is no causal relationship between the premise (i.e., personal attacks preclude a logical objection) and the conclusion (i.e., personal attacks hinder the development of interactions) of this stereotype. The fact that personal attacks are not conducive to reasonable rebuttals originates from the speaker's perspective. However, the role that personal attacks actually play in an interaction depends on the recipient's interpretation and strategic goals. The analysis in Section 5.2.3 has demonstrated that a reply constructed by the speaker with personal attacks is not necessarily understood by the recipient as a personal attack.

The analysis of Interaction 2 also suggests that replies perceived as a personal attack by the recipient can drive the conversation, regardless of their strategic composition by the speaker. For example, in Excerpt 39, User B2's support (Lines 10–11) for @John and User A2 is opposed by User C2 (Lines 12–15).

Excerpt 39 (Interaction 2)

Line	User Account	Text	Level	Turn
01	@John	1R]A desperate cry for help[Crying face][Crying face][Crying face][Crying face]#HurricaneDorian #Abaco #Bahamas Lord please help us [2R][VIDEO: https://twitter.com/i/status/1168227588533293056]	tweet	1
02				
03				
04				
05	User A2	[1R]Father in Heaven, please safeguard Your people in the path of Hurricane Dorian. Loose Your mighty angels to minister unto each of them at this time. Fill them with Your peace that passes all understanding. Keep them together in love. In Jesus' Name, Amen.	1st	2
06				
07				
08				
09				
10	User B2	[1R]And the animals and wildlife too, Lord, created by you.[Folded hands][Folded hands][Folded hands]	2nd	3
11				
12	User C2	[1R]I'm assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn't sound like a very compassionate deity to me.	3rd	4
13				
14				
15				

In response, User B2 (Lines 16–19 & 97–99) takes User C2's objection as an attack and counterattacks User C2 accordingly. However, User B2's response remains overall restrained as a debate with User C2, which indicates User B2's attempt to de-escalate the conflict.

Excerpt 40 (Interaction 2)

Line	User Account	Text	Level	Turn
16	User B2	[1R]Relax Atheist, your day of reckoning will come... everyone's does. But it's funny how, when one's life or a loved one's life is in danger people turn to asking their God to save them. But that would never be you, right??? You'll have to wait and see...	4th	5_1
17				
18				
19				
Omit Lines 18–96				
97	User B2	[1R]And whilst you sit there "assuming", what exactly did you do for the girl in the video who's PLEADING FOR PRAYERS???? If you don't want to pray for her... THEN DON'Tperiod[Neutral face]	4th	5_2
98				
99				

After perceiving this intent, instead of bearing User B2's counterattack in mind, User C2 (Lines 20–23 & 31–37) designs their replies as multiple disagreements. Being voluntarily controlled by Users B2 and C2, the conflictual interaction tends to de-escalate at this point. If not re-intensified, this interaction may gradually come to an end. However, User B2 (Lines 38–41) regards User C2's third 5th level reply (Lines 33–37) as a personal attack and intensifies the conflict by attacking User C2.

Excerpt 41 (Interaction 2)

Line	User Account	Text	Level	Turn
20	User C2	[1R]You are quite correct - it is funny how people to turn to a god to save them from bad things whilst never wondering why their loving all-powerful god allows the things to happen in the first place. The hypocrisy baffles me.	5th	6_1
21				
22				
23				
Omit Lines 24–30				

(To be continued)

Line	User Account	Text	Level	Turn
31	User C2	[1R]Presumably they'll use the all encompassing get out clause that their god moves in mysterious ways.... hmmm....	5th	6_2
32				
33	User C2	[1R]Oh, and by the way - obviously I hope that those that have been affected by the storm are ok - because I'm not a horrible bellend.... I just don't see that praying is likely to help them in any way whatsoever. I was brought up with religion and it has done nothing for me.	5th	6_3
34				
35				
36				
37				
38	User B2	[1R]Then why troll the comments of those that do believe? ... have you nothing better to do, that can help people that are afraid and in danger of dying in a hurricane, who are asking us to pray for them?[1WS] This is your solution?	6th	9
39				
40				
41				

At the 7th level, User D2 (Lines 76–77) joins the interaction and attacks User B2 along with User C2 (Lines 42–46, 66–68 & 71–72) as an atheistic camp.

Excerpt 42 (Interaction 2)

Line	User Account	Text	Level	Turn
42	User C2	[1R]I've not trolled anyone. I've not said anything insulting about any individuals. Just making the point that I fail to see how it will actually help anyone. They are better supported by actual action rather than prayers - whether it be financial or structural aid. Just my opinion.	7th	10_1
43				
44				
45				
46				
Omit Lines 47–65				
66	User C2	[1R]I'd happily donate - and often do - to support those in peril or danger. That surely benefits them more than me telling them I'm thinking of them???	7th	10_2
67				
68				
Omit Lines 69–70				
71	User C2	[1R].... and to be honest - you threatened me with a day of reckoning!	7th	10_3
72				
Omit Lines 73–75				
76	User D2	[1R]They aren't asking you to pray for them. FFS, piss off.	7th	16_1
77	User D2	[1R]Your prayers help no one.	7th	16_2

In order respond to User C2's attacks at the 7th level, User B2 (Lines 47–50 & 73–75) attacks User C2 at the 8th level. Finally, after defending themself against User B2's attack, User C2 (Lines 51–55) gives up attacking. User B2 (Lines 56–60), accordingly, abandons their attack on User C2.

Excerpt 43 (Interaction 2)

Line	User Account	Text	Level	Turn
47	User B2	[1R]Sweetie, these people in the videos where the sorm is now can't get help until it's over. THEY are asking for prayers bc they are scared.[1WS] Do this for them, it's called COMPASSION, and it won't hurt you!!![Folded hands]	8th	11
48				
49				
50				

(To be continued)

Line	User Account	Text	Level	Turn
51 52 53 54 55	User C2	[1R]I can assure you I'm thinking of them, & I'm hoping that they are ok. I just won't pray to a god I don't believe in. I apologise if I caused you offence - I can assure you it wasn't intended that way. I just keep seeing posts about praying, yet bad stuff still happens in droves.	9th	12
56 57 58 59 60	User B2	[1R]Unfortunately the world is made up of good and bad things. We are also given free will to chose a given path bc God didn't make us to be mindless stick figures to experience life in a bubble.[1WS] We make mistakes, and sometimes we need help, or we're thankful.	10th	13_1
Omit Lines 61–72				
73 74 75	User B2	[1R]No. I didn't "threaten" you with a day of reckoning... it's just a fact of life, it happens to everyone one day. You will learn what your honest reactions are when it happens to you..	8th	15

Although User C2's third 5th level reply (see Lines 33–37 in Excerpt 41) is not constructed as a personal attack, it triggers a blatant attack by User B2 at the 6th level (see Lines 38–41 in Excerpt 41): it completely reverses the trend of a premature termination due to conflict management. User B2's attack is perceived by the recipient as an overt provocation because it breaks the relatively stable debate. As a result, this interpersonal conflict escalates into an antagonism between the two camps and eventually culminates at the 7th level (see Excerpt 42). It significantly develops either in terms of the number of participants or the degree of intensity. This reaffirms that it is the recipient's interpretation of the speaker's strategies and their strategic goals that greatly influence the development of a conflictual interaction. In order to explore the impact of personal attacks in conflictual interactions, researchers thus need to focus on the replies that are perceived as a personal attack by the recipient. Such personal attacks may play a critical function in the development of conflictual interactions. They may not only avoid premature termination of an interaction due to control but may also trigger a further intensification of conflictual interactions.

The analysis of personal attacks is also vital for addressing the gaps in CMC studies and linguistic research on interactions. There are no available statistics to prove that personal attacks on CMC are far more frequent and intense than in face-to-face interactions. Personal attacks, however, are indeed a common phenomenon that every CMC user experiences on a daily basis. Users also seem to be empowered with greater freedom to use expressions such as hate speech, attacks and other bald-on-record replies in CMC, to assert their views. Although personal attacks are not unique to CMC, they have become one of the most visible features of CMC. The analysis of personal attacks, thus, becomes crucial to exploring the rules or interactive habits that govern the way people interact in CMC.

However, not only have personal attacks not been studied in detail in linguistics but inharmonious language phenomena have also not received much attention in linguistics; they are regarded as on the fringe of linguistic research. A case in point is Brown and Levinson's (1987/1978) work on linguistic politeness. Their work was produced based on Grice's (1975) cooperative principles, considering the CPs to be the dominant principle that underpins face-to-face interactions and guides interlocutors' actions. Disputes and arguments, in contrast, are seen as impediments to the individuals' interactions. Thus, as mentioned by Brown and Levinson (1987/1978), an important goal of the interaction is to avoid such actions. The approach is what is referred to as linguistic politeness. They argued that linguistic politeness preserves the cooperation between interlocutors, by avoiding or diminishing the harm caused by inharmonious language phenomena. The development of interactions is thus guaranteed. However, the 'cooperative' in the CPs does not refer to the "genuine cooperation or even human benevolence" (Sifianou, 2019, p. 188; see also Leech & Thomas, 1990; Pagliai, 2010; Dynel, 2013). It rather describes the regularities used by people to "interpret utterances which appear untrue, unnecessarily prolix or irrelevant" (Sifianou, 2019, p. 187). Brown and Levinson rather assumed that cooperation leads to harmonious interactions and friendly relationships (Sifianou, 2019). This is also a common issue in current politeness studies. As a result, inharmonious language phenomena (including personal attacks) were for a long time believed to require redress or to be avoided in linguistic research on interactions. Linguists then came to treat them as a reflection of social problems and power imbalances, with less practical application than linguistic politeness. Linguistic research on personal attacks should break the constraints of linguistic conventions on interactions. Personal attacks should not be limited to a linguistic phenomenon that needs to be addressed, redressed or even avoided. They should be taken as an independent linguistic phenomenon that plays as important a role as linguistic politeness in interactions. Further analysis on the relationship between personal attacks and interactions is needed. For example, for what purpose do speakers employ personal attacks? How do personal attacks drive the development of inharmonious interactions? How do they affect the relationship between interlocutors and their relative power? Does technology influence people's use of language on Twitter?

5.4 Summary and conclusion

This chapter provided an overview of how Twitter users employ conflictual linguistic strategies to do conflict in the corpus. As Table 5 suggests, it first discussed the possible reasons for the absence of a de-escalation stage in the corpus. Although Twitter removes posts and suspends accounts in response to users' reports, Twitter's control generally occurs after the termination of conflictual interactions and does not prevent the conflict from intensifying. Thus, conflictual Twitter interactions are unlikely to substantially damage users' reputations, relationships, social status, physical health, etc. in the real world. Twitter users are rarely penalised for observing or engaging in or intensifying conflict. Some of them also regard Twitter as a 'tree hole' to release their repressed emotions in real life by expressing their personal perspectives. As a consequence, users would give little thought to managing or controlling conflict on Twitter. This chapter then explored the strategies for escalating conflictual Twitter interactions in the corpus. To do conflict, Twitter users either object to their opponents' viewpoints (i.e., disagreements) or attack their opponents (i.e., personal attacks) or take both approaches.

Table 5

A summary of findings in Chapter 5

			Previous studies	Chapter 5
Question 1: what are the linguistic strategies employed by Twitter users to engage in conflictual interactions?	De-escalation stage		1. Very common in face-to-face arguments; 2. Functioning to manage/control/ resolve conflict.	1. Not found in the corpus; 2. Maybe due to Twitter users' unique views on face-saving: an in-depth analysis is needed (see Section 7.3 in Chapter 7).
	Escalation stage	Disagreement	1. Very common in face-to-face arguments; 2. Functioning to object to the opponent's view.	1. Very common in conflictual interactions on Twitter; 2. Functioning to object to the opponent's opinion.

(To be continued)

			Previous studies	Chapter 5
Question 1: what are the linguistic strategies employed by Twitter users to engage in conflictual interactions?	Escalation stage	Personal attack	1. Not paying much attention, except for hate speech; 2. Stereotype: personal attacks impede debates.	1. Including hate speech, which is a special type (see Section 2.1.2 in Chapter 2); 2. Common in conflictual Twitter interactions; 3. Functioning to attack the opponent; 4. Seeming to be conducive to the development of conflictual Twitter interactions: an in-depth analysis is needed (see Chapter 6).
		Other	1. Not found in current studies.	1. Not found in the corpus.
Question 2: how do linguistic strategies drive the development of conflictual Twitter interactions?			1. Not paying much attention; 2. Stereotype: divergent views do not foster cooperation.	1. Twitter users take turns to shift their roles from being a recipient of the previous turn to a speaker of the next turn. They interpret the (previous) speaker's turn design, decide their communicative intents and select conflictual linguistic strategies. By doing so, Twitter users cooperate with others and progress conflictual Twitter interactions.

This chapter argued that Twitter users are cooperative with each other in conflictual interactions. This cooperation occurs as a result of Twitter users shifting from being a recipient to being a speaker. Specifically, Twitter users first assume the role of the recipient to interpret the (previous) speaker's turn design (including its use of strategies and communicative intent) and to develop their own strategic goals. Based on these interpretations and goals, Twitter users act as the (next) speaker to construct and post their own turns by using suitable conflictual linguistic strategies and allow their recipients to freely interpret these turns. This suggests that the conversational mechanism used by discourse analysts to explain how people interact in face-to-face settings does not change with changes in the content (i.e., being full of objections and/or attacks) and condition (i.e., being created and constrained by Twitter algorithms) of human communication.

The analysis further found that recipients do not always seem to comprehend or interpret Twitter users' use of strategies and/or strategic goals in the same way as the users; on occasions, it also appears that recipients deliberately avoid constructing the next turns in accordance with their interpretations that reveal Twitter users' communicative intents. This

chapter thus continued to examine this issue and found that both Twitter users' and their recipients' turn designs are of practical significance. These two types of turn constructions both represent their speakers' interpretations of the previous turn and their own strategic goals. Twitter users' turn designs reveal their personal expectations regarding how their turns function in the subsequent conflictual interaction. Their recipients' turn designs are appropriate for indicating the roles that the Twitter users' turns actually play in interactions. This finding suggests that the recipient's actions appear to be more crucial in driving conflictual Twitter interactions than the speaker's use of strategies.

Based on the analysis of Interaction 2, this chapter, finally, discovered that personal attacks appear to play a vital role in progressing conflictual Twitter interactions. Not only may they prevent interactions from being prematurely terminated, but they also greatly increase the intensity of conflict in an evident manner. The speaker's disagreements that are interpreted by the recipient as a personal attack are also likely to serve such functions. However, linguists who study human interactions have not paid much attention to personal attacks. On the one hand, under the influence of the Gricean CPs and politeness theories, linguists have long viewed uninhibited linguistic behaviours, including personal attacks, as the issues to be redressed or avoided. On the other hand, influenced by their preference for applying linguistic theories to society, linguists have also considered addressing these uninhibited linguistic behaviours to be more valuable than investigating their functions in an interaction. Therefore, in the next chapter, this study provides an in-depth analysis of how the replies perceived as a personal attack by the recipient drive conflictual Twitter interactions. The purpose is to continue to further explore the second question, that is, the role of conflictual linguistic strategies in the development of conflictual Twitter interactions.

Chapter 6 The Interactional Function of Personal Attacks

6.0 Introduction

Chapter 5 provided an overview of lingusitic strategies that Twitter users employ to escalate conflictual interactions. It found that the unique features of Twitter interactions created by its algorithms, such as anonymity, no time or geographical restrictions on joining or leaving a conversation and a lack of regulations, challenge the importance of a closing sequence in human communication. It argues that technology appears to be changing people's notions of appropriate language. Chapter 5 also found that to do conflict on Twitter users are likely to object to their opponents' viewpoints (i.e., disagreements) or to attack them (i.e., personal attacks) in a cooperative manner by shifting the role from being a speaker to being a recipient. Both roles are equally important to investigate the developmental mechanism of an interaction: the speaker's turn is suitable for examining how the speaker interprets the turn that they respond to and what role they expect their own turn to play in an interaction. The recipient's turn provides insights into how the (previous) speaker's turn design is interpreted, thus enabling research into the role of the (previous) speaker's turn in the interaction. However, mismatches often occur between the speaker's and the recipient's turns. The analysis of Interaction 2 suggests that such mismatches are particularly evident in interactions in which the recipient perceives that the speaker is to use a personal attack. It was also found that posts that are perceived as a personal attack by the recipient have an important impact on the development of conflictual interactions. Investigating such posts also seems quite useful and practical for developing linguistic research on interactions and addressing the problem of cyber violence.

This chapter, therefore, provides a more comprehensive analysis of personal attacks. It does not identify the function of a post just on the basis of the linguistic strategies that it employs, as is commonly done in research on cyber violence (e.g., Shum & Lee, 2013; Davidson et al., 2017; Decock & Spiessens, 2017b; Mondal et al., 2017; Saleem et al., 2017; Sahlgren et al., 2018; Salminen et al., 2018ab; Siapera et al., 2018; Fernández-Amaya, 2019; Almutairi, 2021). It rather identifies the studied posts serving as a personal attack if they are treated as such by the recipient. This chapter examines in detail the nature of personal attacks, their distribution and how they impact conflictual Twitter interactions. The aim is to investigate the second question (i.e., How do conflictual linguistic strategies drive the

development of conflictual Twitter interactions?) in more depth, building on the findings in Chapter 5 (see Section 5.2) and to explore whether Twitter provides new pragmatic functions for personal attacks.

This chapter starts with making a case for the importance of investigating personal attacks in the corpus and considers its feasibility. Section 6.1 contextualises the focus on personal attacks in this study, examining the distribution and frequency of their occurrence in the corpus. It shows that the analysis of personal attacks can provide comprehensive insights into how Twitter users intensify conflict. Section 6.2 investigates the types of posts in the corpus that are frequently interpreted as a personal attack, their distribution and the factors that determine their occurrence. The results show that the occurrence of personal attacks is highly contextually dependent because there are a lot of personal attacks in the corpus that are constructed in an off-record manner or have the linguistic makeup of disagreements.

The remainder of this chapter subsequently investigates the function of personal attacks from the perspective of the recipient in more detail. It finds that contrary to the stereotype that personal attacks undermine arguments, personal attacks on Twitter function as an acceptable strategy to drive the development of conflictual interactions. Specifically, Section 6.3 is concerned with the impact of personal attacks on the recipient's response. It shows that personal attacks cause irreparable damage to the recipient's face and usually lead to the recipient repairing their own face. The recipient tends to do this by taking on the role of the (next) speaker and issuing a new turn to humiliate the (previous) speaker. Personal attacks thus promote the continuation of conflictual interactions by enhancing the likelihood of receiving a response from the recipient. Section 6.4 shifts the focus to the impact of personal attacks on the nature of conflictual interactions. It investigates the role of personal attacks in different parts of conflictual Twitter interactions. It finds that personal attacks at the beginning of an interaction tend to enhance the aggressive nature of an interaction and to generate opinion camps and viewpoints that remain in place throughout the interaction. They also contribute to broadening the audience of interactions, which is, however, also evident in instances in which personal attacks occur during the course of an interaction. In addition to this, personal attacks in the middle also serve to greatly exacerbate the conflict between interlocutors, trigger a struggle for discursive power between interlocutors and change their relationships. As for personal attacks that occur at

the end of an interaction, in most cases they function to terminate an interaction between the speaker and the attacked person. However, the conflict may reappear in a different interaction involving these two persons or even in a new interaction between that speaker and other persons who hold a similar viewpoint to the attacked person.

6.1 Distribution in the corpus

In order to get a better understanding of conflictual Twitter interactions, this section quantifies the occurrence of personal attacks in the corpus. Chapter 5 found that the linguistic characteristics of personal attacks do not appear to be the only factor that influences how the recipient interprets a post and how this post functions in the interaction. The analysis shows that 5% (185) of the posts (3593) that have the linguistic characteristics of personal attacks are not perceived as an attack by their recipients. This study argues that these posts should not be considered to perform a function as a personal attack as they are not interpreted as such by their recipients.

For ease of understanding, before presenting the results of the analysis, it is necessary to first review two terms that were defined in Section 4.2.1 in Chapter 4: (1) CONVERSATION and (2) stream. A CONVERSATION refers to an interaction, including an original tweet and all posts triggered by that tweet; there are 38 CONVERSATIONS triggered by 38 original tweets in the Twitter Conflictual Interaction Corpus. A stream refers to an one-on-one interaction that derives from an original tweet and ends with no more replies; a total of 5532 streams were found in the corpus.

Table 6 demonstrates that among the 38 CONVERSATIONs, CONVERSATION 26 contains the greatest number of personal attacks; 77% (33) of the posts in this interaction are considered to be a personal attack. No replies in CONVERSATIONs 13 or 15 appear to be interpreted as a personal attack by any of the Twitter users. In this case, the absence of personal attacks is probably due to the limited number of Twitter users who participate in these two CONVERSATIONs. This greatly decreases the possibility of name-calling between interlocutors. Although the distribution of personal attacks varies across CONVERSATIONs, they are prevalent in the corpus and are not concentrated within a few CONVERSATIONs only. They make up 37% (4170) of all the posts in the whole corpus and the proportion of personal attacks in 16 of the CONVERSATIONs exceeds this value (37%).

Table 6

The proportion of the replies taken as a personal attack in the corpus

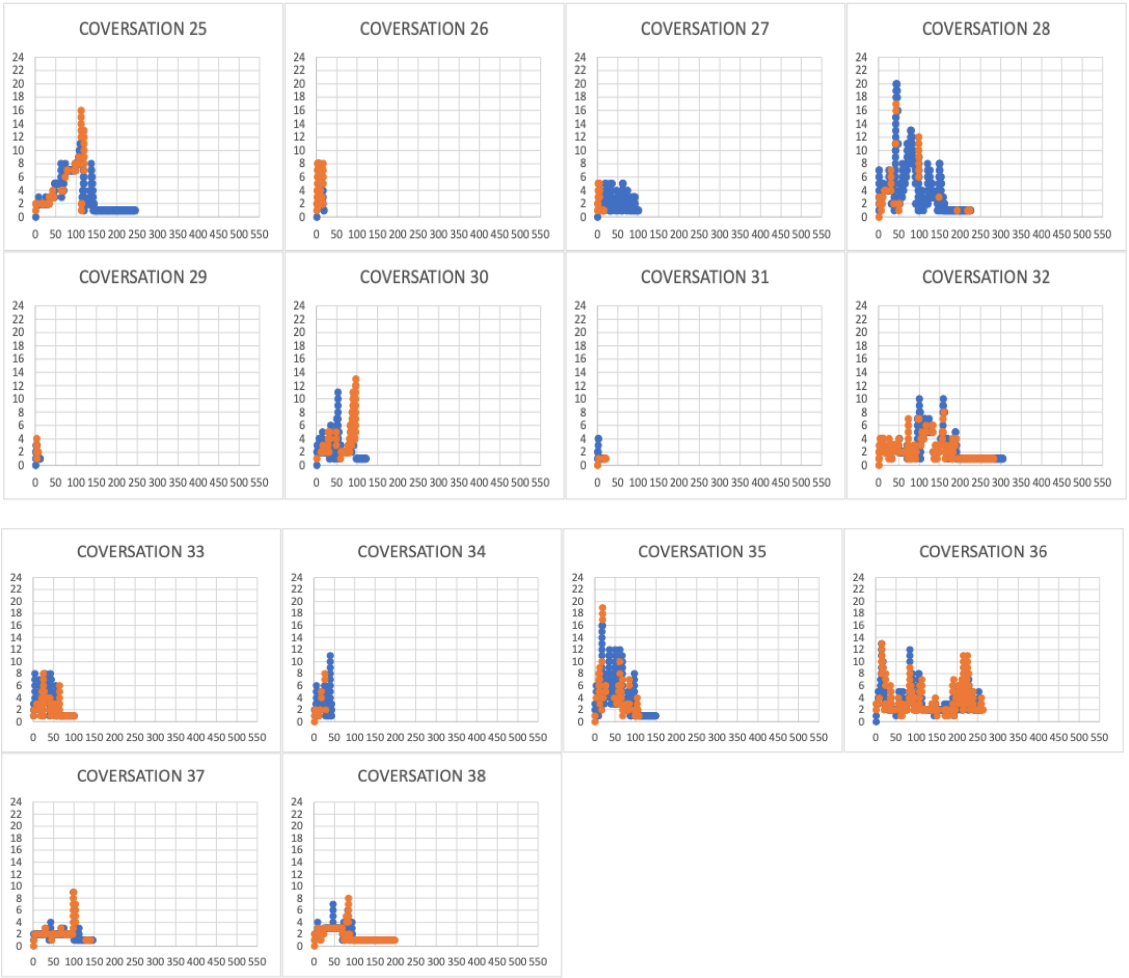
CONVERSATION	Proportion	CONVERSATION	Proportion	CONVERSATION	Proportion
CONVERSATION 1	15%	CONVERSATION 14	30%	CONVERSATION 27	5%
CONVERSATION 2	66%	CONVERSATION 15	0	CONVERSATION 28	9%
CONVERSATION 3	30%	CONVERSATION 16	27%	CONVERSATION 29	12%
CONVERSATION 4	20%	CONVERSATION 17	50%	CONVERSATION 30	30%
CONVERSATION 5	26%	CONVERSATION 18	37%	CONVERSATION 31	23%
CONVERSATION 6	57%	CONVERSATION 19	22%	CONVERSATION 32	55%
CONVERSATION 7	20%	CONVERSATION 20	5%	CONVERSATION 33	44%
CONVERSATION 8	72%	CONVERSATION 21	49%	CONVERSATION 34	10%
CONVERSATION 9	40%	CONVERSATION 22	61%	CONVERSATION 35	21%
CONVERSATION 10	26%	CONVERSATION 23	36%	CONVERSATION 36	45%
CONVERSATION 11	34%	CONVERSATION 24	27%	CONVERSATION 37	27%
CONVERSATION 12	58%	CONVERSATION 25	13%	CONVERSATION 38	58%
CONVERSATION 13	0	CONVERSATION 26	77%	On Average	37%

In order to find out where personal attacks occur within each CONVERSATION, the occurance of personal attacks in each stream in the corpus was examined. The distribution of posts that are taken as a disagreement and those that are regarded as a personal attack in the 38 CONVERSATIONs was plotted on scatter plots in Figure 14. The vertical axis represents 'level' (which indicates the distance from the original tweet); the horizontal axis represents 'stream'; the orange dots represent 'replies using a personal attack', and the blue dots represent 'replies using a disagreement'. It can be seen that in CONVERSATION 2 personal attacks are widely distributed across streams, which is also the case in most CONVERSATIONs in the corpus. However, there are four exceptions to this, which are CONVERSATIONs 13, 15, 25 and 27. According to the previous discussion, the absence of personal attacks in CONVERSATIONs 13 and 15 is very likely due to the limited number of posts in the interaction. The distribution of personal attacks in CONVERSATIONs 25 and 27 is obviously skewed, being mainly concentrated in a few streams. An in-depth analysis of CONVERSATIONs 25 and 27 suggests that these asymmetries are largely attributable to the content of their original tweets. There is no recipient who treats the tweet in CONVERSATION 25, a complaint about the quality of Apple's data cables, as a personal attack. The initial personal attacks in this CONVERSATION emerge at the 1st level. Two 1st level replies are regarded as a personal attack by the next level replies, which leads to the personal attacks in CONVERSATION 25 appearing in only 38 streams out of a total of 245.

The original tweet in CONVERSATION 27 aims to encourage recipients to take part in a survey on the worst make-up products. Although no recipient takes this tweet as a personal attack, 14 streams out of a total of 101 streams contain personal attacks because one 1st level reply is regarded as a personal attack by its recipients. This suggests that personal attacks are also likely to occur in interactions that are triggered by less controversial topics. Once a post is treated as a personal attack, it tends to trigger an accumulation of personal attacks in the subsequent discussion.

Figure 14

The dristribution of replies taken as disagreements and personal attacks in the corpus



The results in this section indicate that it is not a rare phenomenon that Twitter users interpret a post as a personal attack in conflictual interactions. The emergence of personal attacks is also not unique to certain topics or types of interlocutors. This finding preliminarily suggests that personal attacks are an important feature of conflictual Twitter interactions and that the issue of how users interact with each other needs to be examined more closely. The next section, therefore, discusses the factors that condition the emergence of replies that are taken as a personal attack by the recipient.

6.2 Reasons for the recipient feeling attacked and their responses

Although the posts that are designed as containing personal attacks are also interpreted as such in the vast majority of cases in the corpus, there are still 5% (185) of posts that qualify as a personal attack based on their structural features but are perceived as a disagreement. Approximately 14% (577) of the posts that are interpreted as a personal attack by the recipient are constructed by the speaker as merely disagreements. This section, therefore, examines the reasons why the recipient interprets a post as an attack in

more depth employing a qualitative approach. It focuses on CONVERSATION 21 as it epitomises almost all of the cases in the corpus

CONVERSATION 21 starts from a tweet posted on 24th July 2020 by @Jennifer (Lines 01–03), a healthcare worker in the UK, stating her views on mask wearing during Covid-19 (see Excerpt 44). The day before, the NHS had announced that from the 24th July people in supermarkets and shops in England would be required to wear a mask or a face covering. @Jennifer is in agreement with this regulation. She designs her tweet as a combination of a disagreement and a personal attack. In the corpus, this represents the most common type of post that is perceived as a personal attack by the recipient. @Jennifer first implies that mask wearing is not life-threatening by describing her own experience of wearing one, "If I can wear a mask for 13hr on a hot ward" (Line 01). She then advises the public to wear a mask when shopping, "tomorrow you can wear a mask for 20 minutes doing your shopping" (Lines 01–02). However, at the end, she attacks the people who refuse to wear a mask by telling them not to "be a dick" (Line 03).

Excerpt 44

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1

Situation 1: the speaker employs personal attacks

50% (23) of the recipients of @Jennifer's tweet at the first level regard it as a personal attack and attack @Jennifer in response. Most of them are angered by @Jennifer's last phrase, "don't be a dick" (Line 03), and respond with phrases such as "don't be a dick x", "don't be critical of others", "You're the one being a dick!", "but no need to be rude in your tweet" and "don't be offended by someone calling you a dick". Others feel that @Jennifer's suggestion about wearing a mask threatens their negative face. They question @Jennifer's ulterior motive, arguing that she is deliberately spreading fear. They criticise @Jennifer for interfering with their human rights. For instance, "In what universe is it okay to tell everyone else what to do?", "STOP SPREADING FEAR", "shut up", "#GetTheFuckOutOfHere", etc.

Most of the recipients who feel that they are being attacked by @Jennifer construct their replies by employing the same approach as @Jennifer. Not only do they attack @Jennifer but they also object to @Jennifer's view. They insist that wearing a mask is a

personal decision or a government conspiracy or that it does not prevent the spread of the coronavirus. Almost all of these responses are further commented on because these recipients' uses of disagreements are suitable for presenting new viewpoints and stimulating further debates. By doing so, these recipients also prevent conflictual interactions from degenerating into name-calling. A case in point is Excerpt 45. In order to oppose @Jennifer, User A13 argues that mask wearing is a professional or a work-related matter, "you'd have to wear one when u took the job" (Lines 04–05). This leads to a discussion between Users A13 and B13 on the relationship between mask wearing and occupation. User A13 also attacks @Jennifer in an attempt to deprive @Jennifer's right to freedom of expression, "shut up" (Line 05), treating @Jennifer's tweet as an interference with people's freedom. In response, User B13 questions User A13's objection for being pointless and unreasonable, "Make it make sense" (Lines 07–08). They argue that although wearing a mask is a professional requirement for healthcare workers, it does not excuse User A13, who is not a healthcare worker, from not wearing one. Instead of responding directly to User B13's objection, User A13 challenges the validity of @Jennifer's tweet. They argue that even healthcare workers do not wear a mask all the time, "No doctors and nurses have always worn masks" (Line 09). Once again, User B13 (Lines 10–11) undermines User A13's argument as being pointless and ridiculous. User B13 explains that healthcare workers are "Not constantly all day on a ward" (Line 10), and visually expresses their emotions through the emoji of rolling on the floor laughing, "🤣" (Lines 10–11).

Excerpt 45 (Interaction 13)

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
04 05	User A13	[1R]But you knew you'd have to wear one when u took the job so shut up	1st	2
06 07 08	User B13	[1R]She can see into the future? She knew when she took the job there was gonna be a pandemic? Make it make sense lmao	2nd	3
09	User A13	[1R]No doctors and nurses have always worn masks bro	3rd	4
10 11	User B13	[1R]Not constantly all day on a ward they haven't[Rolling on the floor laughing]	4th	5

The remaining recipients at the 1st level who take @Jennifer's tweet as an attack oppose @Jennifer by using very personal attacks. In most cases, their replies do not receive further comments, thus terminating the stream. For example, in Excerpt 46, User A14 suggests that @Jennifer should say the phrase "don't be a dick" (Line 03) in Latin rather than

in English and provides the translation as a quoted image. Although User A14's reply seems to indicate that they are on the same side as @Jennifer, this is not really the case. Since the translation is a word-by-word one rather than a standard Latin expression, User A14 does not really intend to create common ground by providing a more accurate expression. They rather consider Latin, which is nowadays confined to formal and religious settings, to be a symbol of morality and truth. By suggesting a term from Latin, they aim to criticise @Jennifer for pretending to be morally superior and for being pretentious in an off-record manner, thus mitigating the taboo, "dick" (Line 04).

Excerpt 46 (Interaction 14)

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
04 05 06 07	User A14	[1R]I found the Latin for 'do not be a dick' so much more appealing. [2R][PIC: https://twitter.com/IamPaulGibson/status/1287042602143973376/photo/1]	1st	2

Figure 15

The picture in User A14's 1st level reply

Another case in point is Excerpt 47. User A15 criticises @Jennifer's attack as being provocative and pointless, "If I can wear a mask, you can, blah…… Blah…..blah…….blah…… blah……" (Lines 04–05). They then indicate that interacting with @Jennifer is boring through the use of the word "Yawn" (Line 05) and two emoji of sleeping face, "😴" (Lines 05–06). They ultimately challenge @Jennifer's lack of both knowledge and etiquette by arguing that Twitter is a platform without any thresholds, "Every man and his dog say the same thing on here" (Lines 06–07). In doing so, User A15 indirectly attacks @Jennifer for being a troll, thus avoiding a bald-on-record confrontation.

Excerpt 47 (Interaction 15)

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A15	[1R]If I can wear a mask, you can, blah…… Blah…..blah……blah……blah…… [2R]Yawn[Sleeping face][Sleeping face]. Every man and his dog say the same thing on here[Face with tears of joy].	1st	2
05				
06				
07				

It is quite common in the corpus that replies consisting only of personal attacks block conversations. This does not mean, however, that the recipient is incapable or does not feel attacked by the speaker. Instead, it is likely to be the case that the recipient's clear recognition of the speaker's attack prompts them to refrain from responding. No matter how the speaker words their attack, whether off-record (as shown in User A14's Turn 2 in Excerpt 46 and User 15's Turn 2 in Excerpt 47) or bold-on-record (such as "You're either very funny or a stupid cunt" and "Lay off the weed bro"), both types apparently reveal the speaker's intent to attack the recipient. In order to avoid being involved in name-calling, the recipient thus tends to withdraw from the interaction. The reasons for the recipient's withdrawal will be discussed in much more detail in Section 6.4.3.

Situation 2: the speaker does not employ personal attacks

Unlike the aforementioned cases in Situation 1, speakers in this second case merely employ disagreements. Although they do not in any way indicate that they intend to attack the recipient, their responses may still be interpreted by the recipient as a personal attack. The qualitative analysis shows that four factors appear to influence the (mis)interpretation of disagreement as a personal attack.

First, when the speaker's disagreement leaves the recipient with an impression that the speaker takes the moral high ground, this disagreement is likely to be interpreted as a personal attack by the recipient. For example, in Excerpt 48, User A16 opposes @Jennifer by arguing that wearing a mask is a requirement for people in particular occupations, not for everyone, "It's your job" (Line 04). Although User A16 may not intend to attack @Jennifer in response, their reply is treated as a personal attack by User B16 (Lines 05–06). User B16 interprets User A16's reply as an indication that @Jennifer interferes with human rights. They thus criticise User A16 for being selfish, by arguing that wearing a mask is everyone's responsibility to prevent the spread of Covid-19, "it's your job as a decent human being"

(Line 05). From the perspective of User B16, User A16 is hypocritical and in no position to criticise others because User A16 pretends to be a morally superior person.

Excerpt 48 (Interaction 16)

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A16	[1R]It's your job	1st	2
05	User B16	[1R]and it's your job as a decent human being to wear one too :)	2nd	3
06				

Second, if the recipient feels that their right to judge right and wrong has been deprived by the speaker's disagreement, they are likely to consider the speaker's disagreement as a personal attack. It is often the case that disagreements are accompanied by an implicature that indicates the speaker's didacticism or assertiveness. For example, in Excerpt 49, aiming to criticise @Jennifer for lying, User A17 first asks @Jennifer to explain why their personal experience contradicts @Jennifer's statement, "Could you please explain to me why" (Line 04). They then overtly provoke @Jennifer by denying the possibility of receiving an explanation from @Jennifer, "I'm guessing you won't" (Lines 06–07). User B17 counterattacks User A17 by denigrating the professionalism of the hospital that User A17 attended, "You were at sh*t hospital" (Line 08). User A17 thus defends themself by attacking User B17 and @Jennifer as being "insane. Gullible, unthinking" (Line 09).

Excerpt 49 (Interaction 17)

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A17	[1R]Could you please explain to me why when I was in hospital in May for three days most of he doctors and nurses didn't wear masks?[2R:ER] [3R]I'm guessing you won't.	1st	2
05				
06				
07				
08	User B17	[1R]You were at a sh*t hospital	2nd	3
09	User A17	[1R]You people are insane. Gullible, unthinking sheep.	3rd	4

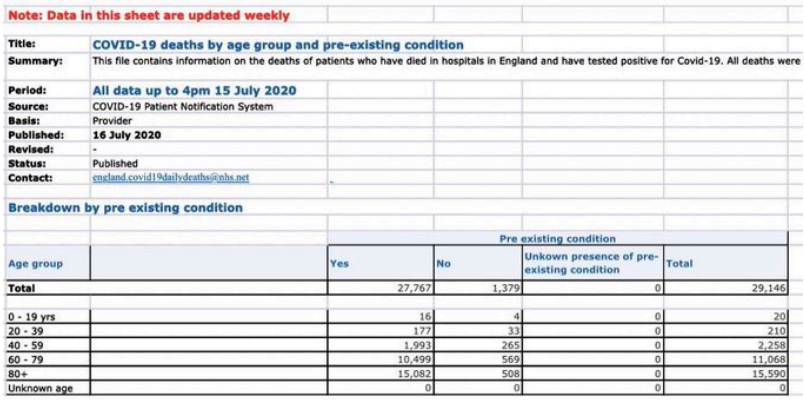
In response, User B17 (Lines 10–14) defends their position, arguing that User A17 lacks discernment and asks them to consult some reliable websites, "Do your own research on RELIABLE websites" (Line 10). In order to reject User B17's accusation, User A17 (Lines 15–17) provides evidence to support their viewpoint, proving that they are able to make sensible judgements. They then attack User B17's level of education, "I presume you can read?" (Line 15).

Excerpt 50 (Interaction 17)

Line	User Account	Text	Level	Turn
10	User B17	[1R]Do your own research on RELIABLE websites like the cdc or any other site with ".org" or ".edu" in the url. Don't rely on ".com" sites like you clearly have been. There's times to stand out from the crowd and there's times to save lives by following others	4th	5
11				
12				
13				
14				
15	User A17	[1R]I do.[1WS] I presume you can read? [2R][PIC: https://twitter.com/Thisaintright12/status/1287109704007680002/photo/1]	5th	6
16				
17				

Figure 16

The picture in User A17's 5th level reply



Note: Data in this sheet are updated weekly

Title:	**COVID-19 deaths by age group and pre-existing condition**
Summary:	This file contains information on the deaths of patients who have died in hospitals in England and have tested positive for Covid-19. All deaths were
Period:	**All data up to 4pm 15 July 2020**
Source:	COVID-19 Patient Notification System
Basis:	Provider
Published:	**16 July 2020**
Revised:	-
Status:	Published
Contact:	england.covid19dailydeaths@nhs.net

Breakdown by pre existing condition

		Pre existing condition			
Age group		**Yes**	**No**	**Unkown presence of pre-existing condition**	**Total**
Total		27,767	1,379	0	29,146
0 - 19 yrs		16	4	0	20
20 - 39		177	33	0	210
40 - 59		1,993	265	0	2,258
60 - 79		10,499	569	0	11,068
80+		15,082	508	0	15,590
Unknown age		0	0	0	0

User A17's (Lines 15–17) attack does not pose a great threat to User B17. User B17 (Lines 18–19) questions User A17's evidence by arguing that their evidence does not support their claim. However, User A17 (Lines 20–21) fails to properly comprehend User B17's challenge. They wrongly assume that User B17 is questioning the authority of their evidence. User A17's clarification, "Works both ways" (Lines 20–21), once again angers User B17. After further clarifying their doubts about User A17's evidence, "I'm telling you that masks work and you just sent me information on patients dying in the uk" (Lines 22–23), User B17 criticises User A17 for being a "bot" (Line 24), which is a "software application that perform automated tasks" on the Internet (Plutis, 2020).[23]

Excerpt 51 (Interaction 17)

Line	User Account	Text	Level	Turn
18	User B17	[1R]You're giving me information on patients dying in the UK?	6th	7
19				
20	User A17	[1R]And you're telling me to check cdc?![2R:ER] [3R]Works both ways[Face with rolling eyes]	7th	8
21				
22	User B17	[1R]I'm telling you that masks work and you just sent me information on patients dying in the uk?? I'm done talking about this bot[Face with tears of joy]	8th	9
23				
24				

[23] Accessed, 22nd February 2022: https://www.avira.com/en/blog/how-to-spot-a-bot-on-social-media

User A17's phrase in Excerpt 51, "Works both ways" (Lines 20–21), conveys their absolute mastery and authority over the truth that cannot be questioned by User B17. This constitutes a serious challenge to User B17's right to judge the truth. From User B17's point of view, User A17 is trying to force User B17 to abandon their own views and knowledge. Hence, User A17's clarification is treated as a personal attack by User B17.

Third, in cases where the speaker's disagreement causes intense annoyance to the recipient, the recipient is also prone to take the speaker's disagreement as a personal attack. For example, in Excerpt 52, User A18 opposes @Jennifer by claiming that there is no scientific basis for the NHS's request on mask wearing, "It was ok to not wear one Thursday but now all of sudden Covid has got worse overnight" (Lines 04–05). After questioning @Jennifer for lying, "You do not wear a mask for 13 hours" (Lines 05–06), User A18 then turns to attack @Jennifer. They tell @Jennifer not to pretend to be morally superior, "but don't critical of others" (Line 08). User B18 responds to User A18's attack on @Jennifer by publicly attacking User A18's selfishness, "I'm definitely going to be critical of people who won't wear a mask for a short amount of time in a shop" (Lines 09–11).

Excerpt 52 (Interaction 18)

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
04 05 06 07 08	User A18	[1R]It was ok to not wear one Thursday but now all of sudden Covid has got worse overnight[Thinking face]You do not wear a mask for 13 hours and if you want to believe magically that wearing a mask 4 months into a pandemic is gona help now then thats yout choice but don't critical of others.	1st	2
09 10 11 12	User B18	[1R]'Don't be critical of others' I'm definitely going to be critical of people who won't wear a mask for a short amount of time in a shop (unless they are exempt), why is it so hard for some people to do I don't understand	2nd	3

Having engaged the interaction with User A18 on whether mask wearing is helpful in five turns, User B18 questions the purpose of User A18 not wearing a mask, "Why are you so against wearing a mask for 20 mins tops??" (Lines 31–32). Although User A18 clarifies their insistence on not wearing a mask, "Cos it wasn't relevant 4 months ago and isn't now" (Line 33), User A18's clarification is still taken as a personal attack by three newly joined Twitter users who had just joined the interaction: Users C18, D18 and E18, see Excerpt 54.

Excerpt 53 (Interaction 18)

Line	User Account	Text	Level	Turn
31 32	User B18	[1R]Why are you so against wearing a mask for 20 mins tops???	8th	9
33	User A18	[1R]Cos it wasn't relevant 4 months ago and isn't now	9th	10_1

User C18 questions User A18's ulterior motives, "just to disrespect them online and potentially cause a second lockdown" (Lines 35–36), and criticises User A18 for being selfish, "someone who won't wear them in public is selfishtwat" (Lines 37–38). User D18 explains that the NHS's request for mask wearing keeps people free and healthy at the same time and criticises User A18 for being an "absolute nonce" (Lines 39–40). User E18 attacks User A18 by arguing that "Shut up you fucking moron" (Line 44).

Excerpt 54 (Interaction 18)

Line	User Account	Text	Level	Turn
34 35 36 37 38	User C18	[1R]I bet you clapped on Thursdays for the NHS doing a great job, just to disrespect them online and potentially cause a second lockdown. You can call ppl who wear masks sheep's all you want, but I think the word for someone who won't wear them in public is selfishtwat.	10th	11
39 40 41 42 43	User D18	[1R]aww fuck up sean and just wear a mask you absolute nonce. As more and more people are now out in public and coming into contact with each other then we need to try reduce the risk of spreading it and a mask does that! Don't want to wear a mask? Then don't go out! Simple!	10th	12
44	User E18	[1R]Shut up you fucking moron	10th	13_1

Many users (including User B18) have explained the reasons and benefits of mask wearing to User A18 several times prior to Turn 10. However, User A18 still continues to insist that wearing a mask is useless and attacks these users from time to time. When User A18's claim about the futility of mask wearing is raised again, Users C18, D18 and E18, therefore, become intensely annoyed with User A18 and treat User A18's clarification in Excerpt 53 as a personal attack. This results in User A18 being perceived as a stubborn, incomprehensible and uncommunicative person.

Fourth, if a key interest of the group to which the recipient belongs or that they favour is interfered with by the speaker's disagreement, the disagreement is likely to be regarded by the recipient as a personal attack. For example, in Excerpt 55, User A19 (Lines 04–05) disagrees with the mandatory mask wearing, arguing that some people are unable to wear one, "some people can't wear them though" (Line 04). As pressed by User B19, "who" (Line 06), User A19 clarifies that they mean people with health conditions, "people with

asthma and other breathing problems, autistic people, some people with anxiety etc" (Lines 07–08). They then continue to criticise the lack of compassion of people who try to make mask wearing obligatory, "obviously if you can where one, you should but some people don't have to" (Lines 08–10).

Excerpt 55 (Interaction 19)

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
04 05	User A19	[1R]some people can't wear them though, "if i can where one, so can you" isn't exactly true	1st	2
06	User B19	[1R]who	2nd	3
07 08 09 10	User A19	[1R]people with asthma and other breathing problems, autistic people, some people with anxiety etc... obviously if you can where one, you should but some people don't have to because they genuinely can't handle it	3rd	4

User C19 opposes User A19 by arguing that wearing a mask does not interfere with breathing, "masks don't stop you from breathing" (Line 12). At the next level, User A19 (Lines 13–16) corrects User C19's opinion and explains the difficulty in mask wearing. They argue that mask wearing makes the aforementioned people breathless, "it's harder to breathe" (Lines 13–14), and creates inconvenience of their lives, "avoid going to shops" (Line 15).

Excerpt 56 (Interaction 19)

Line	User Account	Text	Level	Turn
11 12	User C19	[1R]I know a ton of people with severe asthma, as well as nurses with asthma, masks don't stop you from breathing	4th	5
13 14 15 16	User A19	[1R]they don't stop you but can make you feel like it's harder to breathe, i also know people with asthma that try to avoid going to shops due to being unable to cope wearing one	5th	6

User C19 thus offers a solution of avoiding shopping out, "don't go to the shop" (Lines 17–18), and justifies their view by revealing their identity as an asthmatic, "I have asthma" (Line 18). In response, User A19 goes on to discuss the difficulty in wearing a mask, "but they're that busy atm so sometimes we don't have a choice but to go" (Lines 21–22). To underscore their expertise, User A19 accordingly partially reveals their identity as a patient. They replace the noun "people" (Line 07) with the pronoun "we" (Line 20) for the first time. User A19's statement, however, does not convince User C19 to abandon their objection, "but then you should still put on a mask" (Line 23).

Excerpt 57 (Interaction 19)

Line	User Account	Text	Level	Turn
17	User C19	[1R]I mean, if you really are struggling that much, don't go to the shop. I have asthma and the last thing I'd want is a respiratory infection[Woman shrugging]	6th	7
18				
19				
20	User A19	[1R]that's what we're doing, we keep trying to get deliveries instead but they're that busy atm so sometimes we don't have a choice but to go...	7th	8
21				
22				
23	User C19	[1R]but then you should still put on a mask..	8th	9

At the 9th level, User A19 (Lines 24–26) again tries to assimilate User C19 by revealing their autistic identity and by further explaining the difficulties they meet in wearing a mask. However, User A19's words seriously anger User C19 because what User A19 insists on, i.e., not wearing a mask, poses a serious threat to User C19's life. User C19 attacks User A19 for their selfishness, "stop being selfish" (Line 27), and orders User A19 to not leave home without a mask, "stay at home" (Line 27).

Excerpt 58 (Interaction 19)

Line	User Account	Text	Level	Turn
24	User A19	[1R]i'm autistic and i can't handle it for more than 5 minutes so i cant... i'd much rather wear a mask than risk getting covid but it not that easy..	9th	10
25				
26				
27	User C19	[1R]stop being selfish and stay at home then xoxo cba to argue with people that only care about themselves[Face with rolling eyes]	10th	11
28				
29				

Taken together, the analysis suggests that, in most cases, the use of personal attacks by speakers serves to indicate the occurrence of conflict escalation on Twitter. As Table 7 suggests, if a post consists purely of personal attacks, it is very likely to make the recipient sense an attack from the speaker because of the speaker's apparent desire to do so. In terms of posts that contain both disagreements and attacks, they still have a pretty high possibility of being interpreted as a personal attack, although disagreement parts that are suitable for generating a new argument help to conceal the speaker's intent to attack. However, speakers do not always formulate their attacks in a bald-on-record manner by using apparently abusive and insulting items such as taboo language and dirty words. Even in such cases, the recipient is still able to sense the attack from the speaker. In the corpus, there are also approximately 14% (577) of the posts that are interpreted as a personal attack by the recipient while being constructed by the speaker as mere disagreements. If the recipient believes that they have been deprived of their rights, or the speaker is perceived to take the moral high ground, they are prone to consider the speaker's very objection as an attack. This

interpretation may also be triggered by strong annoyance on the part of the recipient, or if the recipient believes that the speaker attempts to interfere with a key interest of the group to which they belong or which they support. All of these findings suggest that the interactional context strongly influences how users interpret a post on Twitter and thus determines the discursive function of that post. Therefore, from the next section, this chapter starts to explore the interactional function of personal attacks in the conflictual Twitter interactions from two perspectives: (1) their impact on the structure of conflictual interactions (see Section 6.3) and (2) their impact on the nature of conflictual interactions (see Section 6.4).

Table 7

Cases where the speaker's turn is likely to be interpreted as a personal attack

The speaker's design	Being interpreted as a personal attack by the recipient
Personal attacks	1. In most cases
Personal attacks & Disagreements	1. In a lot of cases because of the parts that contain personal attacks;
Disagreements	1. If the recipient believes that they have been deprived of their right to speech; 2. If the recipient believes that the speaker takes the moral high ground; 3. If the recipient believes that the speaker's disagreement is extremely annoying; 4. If the recipient believes that the speaker attempts to interfere with a key interest of the group to which they belong or that they support.

6.3 Function in terms of structure

This section examines the function of personal attacks at the structural level. It investigates how the replies interpreted as a personal attack by the recipient guarantee the continuation of turn-taking and thus maintain the continuation of the interaction. The analysis shows that if the recipient perceives that their face has been severely threatened or damaged by the speaker's strategies (i.e., disagreements, personal attacks, a combination of both), they tend to repair their face through self-defence or counterattacks. The replies that are regarded as a personal attack, therefore, maintain the continuation of the interaction by ensuring a response from the person being attacked.

A case in point is Interaction 2, which was discussed in Section 5.3.1 in Chapter 5. In Excerpt 59, although User C2 (Lines 12–13) does not intend to attack User B2 (i.e., the previous speaker), their criticism of User B2's call for prayer seriously threatens the core beliefs of the theistic community to which User B2 belongs. This forces User B2 (Lines 16–19 & 97–99) to take User C2's objection as a personal attack. To preserve their own face, User B2 provokes and overtly threatens User C2 as an action of self-defence. They argue that User

C2 will one day give up their atheistic belief at the beginning of their reply, “Relax Atheist, your day of reckoning will come...” (Line 16).

Excerpt 59 (Interaction 2)

Line	User Account	Text	Level	Turn
12	User C2	[1R]I’m assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn’t sound like a very compassionate deity to me.	3rd	4
13				
14				
15				
16	User B2	[1R]Relax Atheist, your day of reckoning will come... everyone’s does. But it’s funny how, when one’s life or a loved one’s life is in danger people turn to asking their God to save them. But that would never be you, right??? You’ll have to wait and see...	4th	5_1
17				
18				
19				
Omit Lines 20–96				
97	User B2	[1R]And whilst you sit there “assuming”, what exactly did you do for the girl in the video who’s PLEADING FOR PRAYERS???? If you don’t want to pray for her... THEN DON’Tperiod[Neutral face]	4th	5_2
98				
99				

User B2’s reply, except for the first sentence, generally maintains a restrained style throughout a logically structured rebuttal. In Excerpt 60, User C2 (Lines 20–23 & 31–37), in their response to User B2’s first 4th level reply (Lines 16–19), does not take User B2’s threat seriously. They believe that User B2’s attack on them stems from theist prejudice against atheists, i.e., intolerance towards atheists. User C2 therefore clarifies their view of God’s uselessness, by stating that they also want the affected people to be saved.

Excerpt 60 (Interaction 2)

Line	User Account	Text	Level	Turn
20	User C2	[1R]You are quite correct - it is funny how people to turn to a god to save them from bad things whilst never wondering why their loving all-powerful god allows the things to happen in the first place. The hypocrisy baffles me.	5th	6_1
21				
22				
23				
Omit Lines 27–33				
31	User C2	[1R]Presumably they’ll use the all encompassing get out clause that their god moves in mysterious ways.... hmmm....	5th	6_2
32				
33	User C2	[1R]Oh, and by the way - obviously I hope that those that have been affected by the storm are ok - because I’m not a horrible bellend.... I just don’t see that praying is likely to help them in any way whatsoever. I was brought up with religion and it has done nothing for me.	5th	6_3
34				
35				
36				
37				

However, from the perspective of User B2, User C2’s clarification still implies that God is useless, thus constituting a personal attack. User B2 (Lines 39–41), once again, sets out to preserve their face by blatantly accusing User C2 of being a troll, “why troll the comments of those that do believe?” (Line 38), and denounces User C2 of being unwilling to help. User B2’s overt attack on User C2 totally angers User C2.

Excerpt 61 (Interaction 2)

Line	User Account	Text	Level	Turn
38	User B2	[1R]Then why troll the comments of those that do believe? ... have you nothing better to do, that can help people that are afraid and in danger of dying in a hurricane, who are asking us to pray for them?[1WS] This is your solution?	6th	9
39				
40				
41				

User C2 first defends themself against User B2's overt attack by arguing that "I've not trolled anyone. I've not said anything insulting about any individual" (Lines 42–23) and "I'd happily donate - and often do - to support those in peril or danger" (Lines 66–67). They then explicitly accuse User B2 of committing a personal attack, "you threatened me" (Line 71).

Excerpt 62 (Interaction 2)

Line	User Account	Text	Level	Turn
42	User C2	[1R]I've not trolled anyone. I've not said anything insulting about any individuals. Just making the point that I fail to see how it will actually help anyone. They are better supported by actual action rather than prayers - whether it be financial or structural aid. Just my opinion.	7th	10_1
43				
44				
45				
46				
Omit Lines 47–65				
66	User C2	[1R]I'd happily donate - and often do - to support those in peril or danger. That surely benefits them more than me telling them I'm thinking of them???	7th	10_2
67				
68				
Omit Lines 69–70				
71	User C2	[1R].... and to be honest - you threatened me with a day of reckoning!	7th	10_3
72				

User B2 quotes and rejects User C2's (Lines 71–72) accusation, "I didn't 'threaten' you" (Line 73), and argues that they are only stating a fact.

Excerpt 63 (Interaction 2)

Line	User Account	Text	Level	Turn
73	User B2	[1R]No. I didn't "threaten" you with a day of reckoning... it's just a fact of life, it happens to everyone one day. You will learn what your honest reactions are when it happens to you..	8th	15
74				
75				

The analysis of Interaction 2 suggests that personal attacks arise usually because the speaker's response makes the recipient feel that their face has been damaged unbearably. Instead of hastily withdrawing from the interaction without doing anything to preserve their face, the recipient tends to repair their face by responding to the speaker and by employing personal attacks to construct their response. These personal attacks are also quite likely to be perceived as such by Twitter users who respond to them because Twitter users' personal attacks in 95% (3408) of cases in the corpus are felt to seriously humiliate their opponents. Conflictual interactions continue to develop as a result of this cycle in which attack-inspired

defences are transformed into new attacks. However, this is not absolute. Approximately 5% (185) of the replies in the corpus that include personal attacks are not taken as such by the recipient at all or are even left unanswered. For instance, as shown in Excerpt 63, User B2's attack (Lines 73–75) at the 8th level on User C2 is no longer commented on. This phenomenon was also noticed in Section 6.2, when analysing the replies that consist entirely of personal attacks and are regarded as such. The streams in which these replies are located may terminate following an attack, but the interaction between the interlocutors may subsequently continue elsewhere in another stream. For example, although User B2's attack (Lines 73–75) at the 8th level was not responded to by User C2, their interaction still continues. This phenomenon will be discussed in more detail in the next section (see Section 6.4.3), since it relates to the function of the replies that are viewed as a personal attack by the recipient and are located at the end of an interaction.

As Chapter 5 suggests that conflictual Twitter interactions lack a closing sequence (see Section 5.1.1), the analysis of personal attacks here also finds that the recipient sometimes voluntarily terminates the interaction by withdrawing from it, when sensing that they have been attacked by the speaker. In addition to the technical reasons outlined in Chapter 5, it is also possible that, from a pragmatic perspective, the recipient withdraws from the debate because they feel that their response is not helpful to the argument or because they do not want to expose themself to a name-calling war. Although in such cases personal attacks may prevent the continuation of conflictual interactions, they do intensify the conflict between interlocutors in terms of the overall trend of interactions. For example, in Excerpt 64, @washingtonpost (Lines 01–09) posts a tweet, reporting the death of Qasem Soleimani. Although this tweet is designed as a news report, its use of the phrase "Iran's most revered military leader" (Lines 01–02) receives a lot of criticisms in CONVERSATION 6. User A20's Turn 2 is a case in point. To oppose @washingtonpost, User A20 claims that Qasem Soleimani is a "legit Islamic terrorist" (Line 10), and argues that @washingtonpost, as an American news outlet, should support this stance. Following an empty line, "[2R:ER]" (Line 10), they then escalate the opposition and publicly accuse @washingtonpost of betraying the American people, "you guys aren't even trying hide your hatred of America anymore" (Lines 10–11). This results in the subsequent conflictual interaction between Users B20 and C20 from Turn 3 to Turn 8.

Excerpt 64 (Interaction 20)

Line	User Account	Text	Level	Turn
01	@washingtonpost	[1R]Breaking news: Airstrike at Baghdad airport kills Iran's most revered military leader, Qasem Soleimani, Iraqi state television reports [2R][QUOTE: https://www.washingtonpost.com/world/national-security/defense-secretary-says-iran-and-its-proxies-may-be-planning-fresh-attacks-on-us-personnel-in-iraq/2020/01/02/53b63f00-2d89-11ea-bcb3-ac6482c4a92f_story.html?utm_campaign=wp_main&utm_medium=social&utm_source=twitter]	tweet	1
02				
03				
04				
05				
06				
07				
08				
09				
10	User A20	[1R]He was a legit Islamic terrorist.[2R:ER][3R]My God, you guys aren't even trying hide your hatred of America anymore.	1st	2
11				

User B20 (Lines 12–16) objects to User A20's use of the phrase "legit Islamic terrorist" (Line 10). They distinguish Qasem Soleimani from "OBL" (Osama Bin Laden) and "Al-Baghdadi" (Line 14), arguing that Qasem Soleimani is "not an 'international terrorist'" (Line 12) but a "high ranking military officer of [Iran]" (Lines 12–13), thus supporting @washingtonpost. This argument is further illustrated through an analogy between the death of Qasem Soleimani and that of "[the] Chairman of the Joint Chiefs [of the US]" (Lines 14–15). User B20 finally questions User A20's viewpoint as nationalistic, "spout stupid MEGA taking points" (Lines 15–16). MEGA refers to a political slogan used by Donald Trump, 'make America great again' (Dictionary—MEGA, 2022). In the meantime, User B20 also attempts to avoid attacks from other Americans by using the word "our" (Line 14) to identify themself as a citizen of the United States. However, User C20 still views User B20 as an anti-American person and overtly challenges User B20's identity as a pro-democracy activist, a political consultant and an Iraq War veteran, which is displayed on their profile page and has been verified by Twitter, "who did you blow to get a blue check mark?" (Line 17). The blue check mark is a symbol used by the Twitter company to signify that "an account of public interest is authentic" ("Twitter—About Verified Accounts", 2022).[24] This marks the beginning of the escalation from an opinion argument to a personal attack.

Excerpt 65 (Interaction 20)

Line	User Account	Text	Level	Turn
12	User B20	[1R]He was not an "international terrorist". He was a high ranking military officer of a nation. That's not the same as OBL or Al-Baghdadi. This is like Iran taking out our Chairman of the Joint Chiefs. Do you know anything or just spout stupid MAGA taking points?	2nd	3
13				
14				
15				
16				
17	User C20	[1R]who did you blow to get a blue check mark?	3rd	4

[24] Accessed, 22nd February 2022: https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts

User B20, in return, challenges that the blue check mark of User C20's account was obtained in an improper way, "Wait that's how you got yours?" (Line 18). After clarifying their identity and trustworthiness, User B20 criticises that User C20 has ulterior motives, "Who did you have to do?" (Lines 19–20). User C20 responds to User B20's attack with a provocation, arguing that they got the verification by "storm[ming] Twitter headquarters" (Line 21). This leads to User B20 once again criticising User C20 for having ulterior motives, "not aware of that as a technique" (Line 22). They then challenge User C20 for being a bot and ask User C20 to demonstrate that they are "a human" (Line 23). By replying with the phrase "me too" (Line 24), User C20 retaliates against User B20, attacking User B20 for being a bot off-record and reinforces their attack through the use of "hon" (Line 24). After that, User B20 no longer responds to User C20, making User C20's attack the final turn.

Excerpt 66 (Interaction 20)

Line	User Account	Text	Level	Turn
18	User B20	[1R]Wait that's how you got yours? I thought it was my decades of military service and sparkling personality. Who did you have to do?	4th	5
19				
20				
21	User C20	[1R]storm Twitter headquarters in San Francisco.	5th	6
22	User B20	[1R]Again...was not aware of that as a technique. I did it when you still had to apply and prove you were a human.	6th	7
23				
24	User C20	[1R]me too hon	7th	8

Although user C20's attack in Turn 8 terminates the interaction, as for the overall structure, this attack functions as a continuation and escalation of their attack in Turn 4. Together, these turns successfully escalate Interaction 20 from a debate about the wording used by @washingtonpost to an unjustified attack on the interlocutor's characteristics. Even though personal attacks may prevent the continuation of a conflictual interaction, they can also contribute to the overall development of the interaction. This is because personal attacks in this case appear to alter the nature of conflictual interactions and to bring the interaction into a new stage, which will be discussed in more depth in the next section.

Overall, the discussion in this section suggests that personal attacks do play a key role in the development of conflictual Twitter interactions at the structural level. The analysis identified two ways in which this happens. Personal attacks in many cases receive responses, leading to the continuation of interactions because of their potentially face-threatening nature. If the recipient interprets a reply as a personal attack, this reply also automatically poses a damage to the recipient's face as the two are intimately linked. It is the latter that motivates the recipient's response, which constitutes an attempt at saving face. The

recipient shifts their role from being a recipient to being a (next) speaker and designs their response as an attack on the previous speaker. There are also cases where interactions apparently terminate after the recipient feels attacked. In the majority of cases, interactions do end, however. Even in these cases they also contribute to the development of conflictual interactions because they escalate the conflict. In other cases, the termination is temporary. Interactions continue in different streams that are fuelled by the personal attacks in the previous streams. The next section continues to examine the impact of personal attacks on the nature of conflictual interactions.

6.4 Function in terms of nature

This section explores the impact of the placement of personal attacks within turns on the nature of conflictual Twitter interactions. This issue is addressed by examining the role of personal attacks located at the beginning, in the middle and at the end of conflictual Twitter interactions, respectively.

6.4.1 Beginning

As each interaction in the corpus starts from a tweet, personal attacks that occur at the beginning of a conflictual Twitter interaction refer to the original tweets that are treated as a personal attack by their recipients at the 1st level. Of the 38 original tweets in the corpus, 24 or over 63%, are taken as a personal attack by more than one recipient. The purpose of this sub-section is to examine whether the tweets that are regarded as a personal attack have an impact on the scale, sentiment and topics of the entire CONVERSATIONs they elicit.

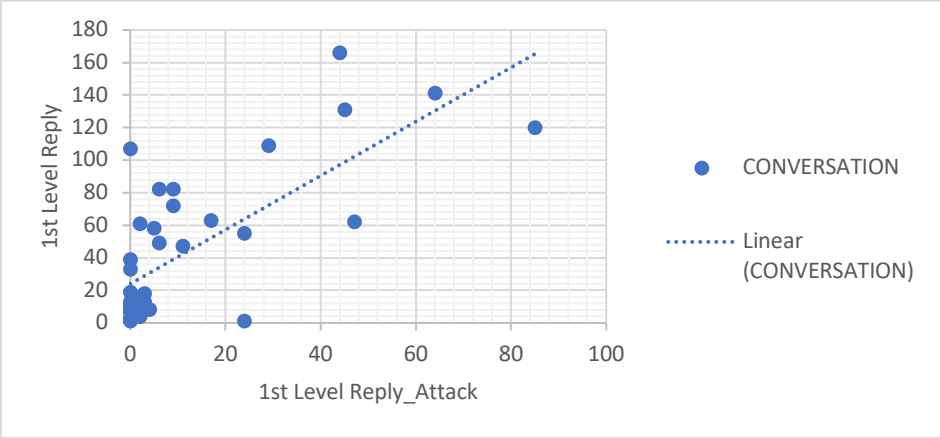
Impact 1: scale and sentiment

In order to examine the impact of personal attacks on the scale and sentiment of conflictual interactions, the Pearson's *r* was calculated. Specifically, it was used to analyse whether tweets that are more likely to be taken as a personal attack: (1) attract more Twitter users to engage in direct arguments with the tweets, (2) generate more discussions in the CONVERSATION, (3) lead to longer one-on-one interactions and (4) result in the COVERSATION becoming more aggressive. The results are as follows:

First of all, although Twitter users are able to respond to the original tweet with an unlimited number of 1st level replies, they only post one reply in most cases. The number of Twitter users who directly argue with the author of the tweet is, in general, reflected by the total number of the 1st level replies to that tweet. As a result, the number of the 1st level replies taking the tweet as a personal attack has a highly positive correlation (r≈0.83) with the total number of the 1st level replies in the CONVERSATION triggered by that tweet. This demonstrates that as Figure 17 shows, the more likely a tweet is to be perceived as a personal attack, the more likely it is to provoke more Twitter users to engage in a direct argument with it.

Figure 17

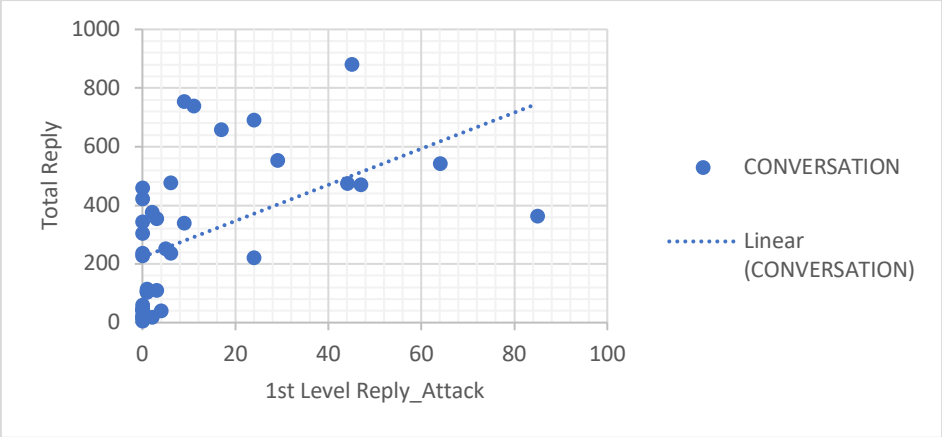
1st level replies taking a tweet as a personal attack vs. 1st level replies in the CONVERSATION triggered by the tweet



Second, the amount of discussion in a CONVERSATION can be gauged by the total number of its replies. The result shows a relatively strong positive correlation (r≈0.51) between the number of the 1st level replies taking a tweet as a personal attack and the total number of replies in the CONVERSATION triggered by that tweet. This suggests that as Figure 18 shows, the more likely a tweet is to be perceived as a personal attack, the more likely the CONVERSATION that stems from that tweet is to receive discussions.

Figure 18

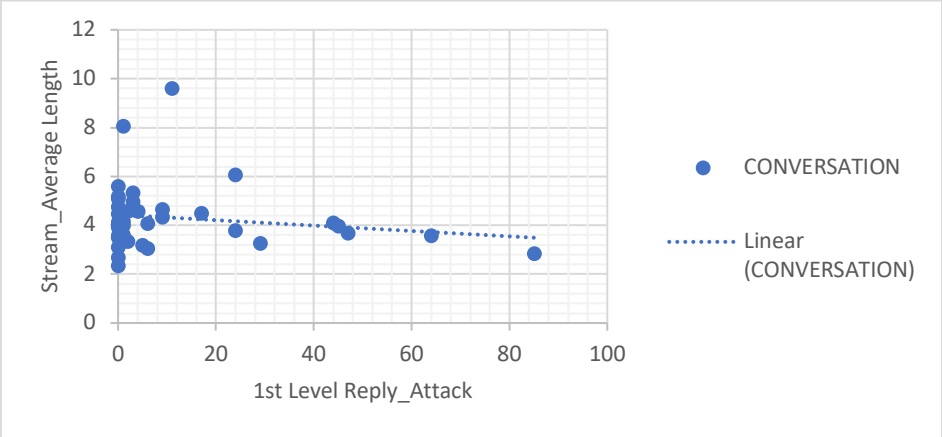
1st level replies regarding a tweet as a personal attack vs. replies in the CONVERSATION triggered by the tweet



Third, the correlation between the number of 1st level replies regarding the tweet as a personal attack and the average level included in streams in the CONVERSATION triggered by that tweet was assessed.[25] The levels contained in an interaction reflect the length of one-on-one interactions in that interaction, i.e., the length of streams. However, no obvious correlation between the two variables was identified (r≈-0.16). This suggests that as Figure 19 shows, whether or not a tweet is perceived as a personal attack by more 1st level replies has little impact on the average length of one-on-one interactions in the CONVERSATION triggered by that tweet.

Figure 19

1st level replies regarding a tweet as a personal attack vs. the average length of streams in the CONVERSATION triggered by the tweet

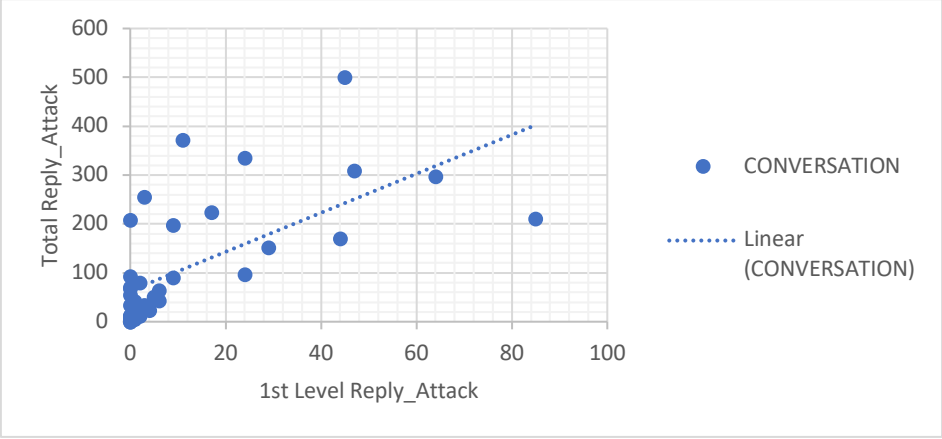


[25] Stream refers to one-on-one interactions that originate from a tweet and end with no more replies (see Section 4.2.1 in Chapter 4).

Fourth, calculating the sentiment scores of CONVERSATIONs is a suitable and general method of analysing their degree of aggressiveness. This is because sentiment scores are used to identify the polarity of information, classifying it into categories such as positive, negative and neutral (McEnery & Hardie, 2011). However, they are not suitable for this study because aggressiveness in this study is measured by the recipient's response, which differs from the polarity of words or sentences. Measuring the aggressiveness of CONVERSATIONs through the number of their replies that are treated as a personal attack by at least one recipient becomes the most appropriate approach. The results show the correlation between the number of 1st level replies perceiving the original tweet as a personal attack and the number of replies regarded as a personal attack in the CONVERSATION triggered by that tweet gives a value around 0.64. This indicates that as Figure 20 shows, the more likely a tweet is to be perceived as a personal attack, the more offensive the CONVERSATION generated from that tweet is likely to be (see Figure 20).

Figure 20

1st level replies regarding a tweet as a personal attack vs. replies treated as a personal attack in the CONVERSATION triggered by the tweet



Impact 2: topics

The focus, subsequently, turns to the impact of tweets that are treated as a personal attack on the topics of the CONVERSATION they generate. This issue was addressed by comparing the differences in the occurrence of opinion camps between CONVERSATIONs triggered by the tweets that are commonly seen as a personal attack and those elicited by the tweets that are never perceived as a personal attack.

On average, about 27% (443) of 1st level replies in the corpus regard their tweets as a personal attack. Tweets in 9 CONVERSATIONs are treated as a personal attack by more than 27% of their 1st level replies. The interactional topics in these 9 CONVERSATIONs generally emerge at the 1st level and, as a result, trigger relatively stable opinion camps. A possible explanation for this might be that tweets which can easily be interpreted as a personal attack often carry a single topic. This leads to their conversations being highly transparent. Twitter users can easily identify the topic and opinion camps and are able to easily position their views, thus giving rise to a strongly ordered or fixed interaction in which all participants focus their attention on attacking their opponents. Therefore, in the 9 CONVERSATIONs, the discursive topics and opinion camps of participants are clear and relatively fixed from beginning to end. The fixity of topics and opinions also further results in a high level of aggression in these 9 CONVERSATIONs. For instance, CONVERSATION 2 begins with @Paul's tweet (see Excerpt 67) about his heartbreak over the Amazon rainforest fires and his desire to lend a helping hand. However, this tweet, which is apparently non-offensive, is regarded as a personal attack by more than 76% (47) of its 1st level replies because its recipients link the tweet with @Paul's previous black history of mocking the dead in Aokigahara forest in Japan.[26] They had previously formed a stereotype about @Paul's disrespectful behaviour towards the dead and nature and now transfer it and their resentment to @Paul's reaction to the Amazon fire. No matter what @Paul says in his tweet, he is still generally seen as acting selfishly against people and nature.

Excerpt 67

Line	User Account	Text	Level	Turn
01 02 03 04 05 06 07	@Paul	[1R]amazon rainforest: how can i help? need advice on how i can actually make a change here; these pictures are breaking my heart... this is one of the most important ecosystems on earth. is there a fundraiser, a call to action, anything i can leverage my audience for on this? [2R][PIC: https://twitter.com/LoganPaul/status/1164064145198600193/photo/1]	tweet	1

[26] While less than 25% (14) of the remaining 1st level replies are in support of @Paul, these replies are retained in the corpus as they are part of the conflictual interactions that are generated next.

Figure 21

The picture in @Paul's tweet

A case in point is Interaction 21 shown in Excerpt 68. User A21 attacks @Paul by implying that @Paul is a plague to the rainforest, "all i seen was logan paul and rainforest and my heart sank" (Line 08). User A21's act is rationalised as non-provocative, "no hate btw" (Line 09), thus further emphasising the attack. In response, User B21 (Lines 10–11) publicly denounces User A21's attack on @LoganPaul, "stop criticising" (Line 10), and continues with an overt attack. They criticise User A21 as a bot who deliberately picks fights with others without offering substantial assistance, "you dont bring anything to the table" (Lines 10–11). In the last turn of this interaction, User C21 enters the conversation and implies that User B21 is the one who is a bot, "And you did?" (Line 12). The analysis suggests that the topic of Interaction 21 is sparked by @Paul's tweet about whether or not @Paul's previous behaviour should be forgiven. This topic is also found in other interactions in CONVERSATION 2 and is maintained throughout the CONVERSATION from the moment it appeared at the 1st level. Based on this topic, participants in CONVERSATION 2 are naturally divided into two groups: (1) a pro-@Paul group and (2) an anti-@Paul group. This is a division which also lasts until the end of the CONVERSATION.

Excerpt 68 (Interaction 21)

Line	User Account	Text	Level	Turn
08	User A21	[1R]all i seen was logan paul and rainforest and my heart sank	1st	2_1
09	User A21	[1R]no hate btw[Folded hands][Black heart]	2nd	2_2
10	User B21	[1R]Mate stop criticising when even you dont bring anything to the table. So please just get a life.	2nd	3
11				
12	User C21	[1R]And you did?	3rd	4

There are also 14 CONVERSATIONs in the corpus whose replies are not taken as a personal attack. This can be attributed to three main reasons. First, some of the 14 tweets are daily banter, complaints or jokes about the speaker's experience around non-sensitive

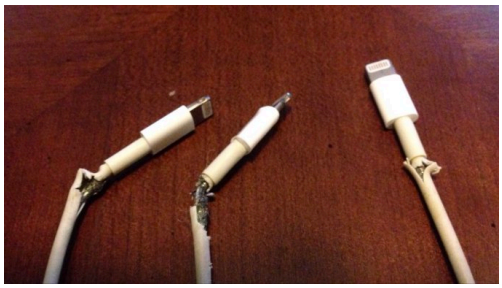
topics. They do not generally pose a threat to the majority of the recipients. They rather evoke a collective memory of the subject of the tweet, which in turn triggers collective complaints on the topic, either in opposition to or in approval of the viewpoints of the tweet. As a result, Twitter users barely take this type of tweet as a personal attack. For example, the tweet of CONVERSATION 25 shown in Excerpt 69 satirises the fact that Apple, a company with a high market capitalisation, produces data cables of poor quality. This is a lived experience for most Twitter users. Of all the 1st level replies that have triggered conflictual interactions, some are shocked at how this cable could be used in such a way; some argue that the quality of Apple's cables is irrelevant to the company's market value, but none take this tweet as a personal attack.

Excerpt 69

Line	User Account	Text	Level	Turn
01 02 03 04	@Nancy	[1R]As a trillion dollar company, you'd think that #Apple could figure out a way to make better cords. [2R][PIC: https://twitter.com/NotDonnaBrazile/status/1025051419907379201/photo/1]	tweet	1

Figure 22

The picture in @Nancy's tweet



Second, some tweets in the 14 CONVERSATIONs only intend to report news or to disseminate information in a broadcast tone. They are hardly ever interpreted as a personal attack because they contain no obvious statement of opinions. They are often used as a source of information to provide a topic for further discussion. For example, the tweet in CONVERSATION 1 shown in Excerpt 70 reports Notre-Dame's fire. All 1st level replies pray for Notre Dame or inquire about the cause of the fire, but none of them object to @BBCBreaking's tweet or take it as a personal attack.

Excerpt 70

Line	User Account	Text	Level	Turn
01 02	@BBCBreaking	[1R]Structure of Notre-Dame saved from total destruction, Paris fire official says https://bbc.in/2UAMlVn	tweet	1

Third, some tweets seen in the 14 CONVERSATIONs are rarely considered as a personal attack because they provide a topic for subsequent interactions and a venue for these interactions, rather than providing a perspective. For example, the tweet in CONVERSATION 27 displayed in Excerpt 71 aims to invite Twitter users to discuss make-up products without any opinion output, so none of the 1st level replies oppose this tweet or treat it as a personal attack.

Excerpt 71

Line	User Account	Text	Level	Turn
01 02	@Kelly	[1R]What are THE WORST makeup products you've tried? It's for science.	tweet	1

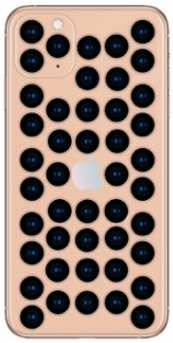
The analysis shows that the topics in these 14 CONVERSATIONs are not as focused as those in those 9 CONVERSATIONS that are used to respond to the tweets that are generally perceived as a personal attack. The topics in these 14 CONVERSATIONs are inspired, rather than carried, by the tweet; they are also hardly sustained from the 1st level reply to the end of the interaction as they are prone to change significantly in the subsequent interaction. This also leads to the emergence of multiple opinion camps in the interaction, with participants being split up. Twitter users are also less aware of which side of the discussion they are going to participate in and support because they read the tweets without knowing what camps of opinion are coming up in the ensuing interaction. The overall aggressiveness of the 14 CONVERSATIONs triggered by such non-aggressive tweets is, therefore, significantly reduced. However, this is not to suggest that these tweets cannot trigger highly conflictual Twitter interactions. It is just that personal attacks that emerge in the oncoming interaction are neither an attack on the speaker of the original tweets nor are they directly related to the views expressed in those tweets. For instance, over 77% (33) of the replies in CONVERSATION 26 treat @David's tweet as a personal attack and are all triggered by one 1st level reply from User A22. In the tweet, @David makes fun of iPhone's increasing lack of design, "#iphone 15 - It's All Cameras!" (Line 01). In response, User A22 disapproves of @David's sense of humour, saying that @David's photo makes them feel uncomfortable, "takes me awhile to get over seeing something like this" (Lines 05–06).

Excerpt 72 (Interaction 22)

Line	User Account	Text	Level	Turn
01	@David	[1R]coming soon: #iphone 15 - It's All Cameras!™ #AppleEvent [2R][PIC: https://twitter.com/starboots_/status/1171490143942758401/photo/1]	tweet	1
02				
03				
04				
05	User A22	[1R]People think this is funny, but it seriously takes me awhile to get over seeing something like this.	1st	2
06				

Figure 23

The picture in @David's tweet



However, out of a total of seven recipients who raise an objection, five take User A22's reply as a personal attack, who are Users B22 (Line 07), C22 (Line 08), D22 (Lines 09–10), E22 (Lines 11–12) and F22 (Lines 13–14). They question whether User A22 is deliberately drawing attention. Users B22 and C22 also insult User A22 as a "fucking ape" (Line 07) and a "clown" (Line 08). Twitter users who participate in the remainder of the interaction are, accordingly, divided into two camps based on whether or not they justified the hatred of the circle to be grandstand. They present the entire replies that are taken as a personal attack in this CONVERSATION, suggesting that the attacks here are no longer due to the complaints about iPhone's design in @David's original tweet.

Excerpt 73 (Interaction 22)

Line	User Account	Text	Level	Turn
07	User B22	[1R]are you a fucking ape?	2nd	3
08	User C22	[1R]so don't look at it[Clown face]	2nd	4
09	User D22	[1R]Imagine being so attention hungry u tell other people and tell urself ur scared of holes lol	2nd	5
10				
11	User E22	[1R]People aren't gonna feel sorry for you because you got in your feels over a post on twitter	2nd	6
12				
13	User F22	[1R]Don't be a fucking weakling then. Take control over your body	2nd	7
14				

The comparison of the two results suggests that tweets that can easily be interpreted as a personal attack often encompass a topic that tends to motivate their audiences to

evaluate/judge the author of the tweet. Conflictual Twitter interactions triggered by these tweets are thus likely to produce a fixed and persistent topic, as well as an opinion camp. The concentration of topics and participants, in turn, enhances the overall aggressiveness of the interaction.

6.4.2 Middle

Personal attacks that occur in the middle of a conflictual interaction refer to replies that are used in response to a previous reply and also receive further responses, so they are generally located in a conflictual Twitter interaction between the 1st level and the penultimate level. The analysis suggests that such personal attacks influence and drive the development of subsequent interactions in the following four ways: (1) significantly intensifying the conflict, (2) changing the purpose of doing the conflict, (3) reforming the relationship between interlocutors and (4) extending the reach of interactions (in given situations).

Function 1: significantly intensifying the conflict

Replies that are taken as a personal attack by the recipient generally serve to exacerbate the aggression between interlocutors regardless of their positions in the interaction. For example, as discussed in Section 6.3, in Excerpt 74, User C20's personal attack (Line 24) on User B20 (Lines 22–23) further intensifies the conflict, by criticising User B20 for being a bot in an off-record manner and reinforcing their attack through the use of the word "hon" (Line 24), and terminates the interaction.

Excerpt 74 (Interaction 20)

Line	User Account	Text	Level	Turn
22 23	User B20	[1R]Again...was not aware of that as a technique. I did it when you still had to apply and prove you were a human.	6th	7
24	User C20	[1R]me too hon	7th	8

However, what this section focuses on are replies that are located in the middle of a conflictual interaction, consist entirely of personal attacks and are also perceived by the recipient as an attack. Since these replies do not help to deepen the discussion on a particular topic and also often lead to name-calling, personal attacks tend to be stereotypically regarded as detrimental to the development of an argument, but this stereotype is not confirmed by the analysis. A case in point is Interaction 23 that is triggered

by @Jennifer's tweet (Lines 01–03) in Excerpt 75 (see Appendix 1 for the full version). @Jennifer tweets about her opinion on mask wearing, which is, however, taken by User A23 as a personal attack. User A23 opposes @Jennifer by arguing that wearing a mask is not helpful at all, "with no PPE. Guess what.. I'm still here" (Line 05). They continue to counterattack @Jennifer by ordering her to stop spreading hate, "STOP SPREADING FEAR" (Line 06). In response, User B23 mimics User A23's approach and attacks User A23 by restricting User A23's right to speak, "Shut it karen" (Line 07). User A23 then escalates their attack by criticising User B23's character for being callous, "You're hard" (Line 08).

Excerpt 75 (Interaction 23)

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
04 05 06	User A23	[1R]Oh I was also walked right through a 'red zone' accidentally with no PPE.[1WS] Guess what.. I'm still here. [2R:ER] [3R]STOP SPREADING FEAR.	1st	2
07	User B23	[1R]Shut it karen	2nd	3
08	User A23	[1R]You're hard.[Face with rolling eyes]	3rd	4

In response to the attack on their personality, User B23 further intensifies the conflict by attacking User A23's intelligence, "but you be an idiot" (Line 10). Accordingly, User A23 attacks User B23, insulting User B23's intelligence and questioning User B23's level of education, "Can you spell potato?" (Line 12). This is endorsed by User C23 at the 6th level, "They still wont get it" (Line 15), thus forming a camp against User B23 in this interaction. Being further angered, User B23 defends themself by overtly criticising Users A23 and C23 as "minging" persons (Line 16).

Excerpt 76 (Interaction 23)

Line	User Account	Text	Level	Turn
09 10	User B23	[1R]You're so solid that you won't wear a piece of cloth to help the cause but you be an idiot	4th	5
11 12 13 14	User A23	[1R]'Help the cause'[Face with tears of joy][2R:ER] [3R]Good one.[1WS] Can you spell potato? [4R][PIC: https://twitter.com/Thisaintright12/status/1287031639323881474/photo/1]	5th	6
15	User C23	[1R]They still wont get it[Face with tears of joy]	6th	7
16	User B23	[1R]Your kids minging	7th	8

Figure 24

The picture in User A23's 5th level reply

At the next two levels, User D23 (Lines 17–19) replaces User A23 and escalates the attack on User B23 by insulting the body parts of User B23 and their relatives, "a cum stain on your dads boxers pal" (Line 17) and "Little billy shrimp dick" (Line 18). In response, User B23 (Lines 20–21) criticises all of their opponents, including @Jennifer, for being awkward throughout the interaction. User B23 uses the phrase "cry more" (Line 20) commonly used in online games and pairs the phrase with a GIF showing a specific action. User A23 (Lines 22–23) returns to the interaction and comments on User B23's performance. User A23 criticises User B23 for behaving like a "keyboard warrior" (Line 22).

Excerpt 77 (Interaction 23)

Line	User Account	Text	Level	Turn
17	User D23	[1R]If only you were still a cum stain on your dads boxers pal	8th	9_1
18	User D23	[1R]Little billy shrimp dick verbally attacking a kid[Face with tears of joy]	9th	9_2
19				
20	User B23	[1R]Cry more [2R][GIF: https://twitter.com/i/status/1287064860434538499]	10th	10
21				
22	User A23	[1R]What a sad little keyboard warrior you are. [Face with tears of joy]	11th	11
23				

Figure 25

The GIF in User B23's 10th level reply

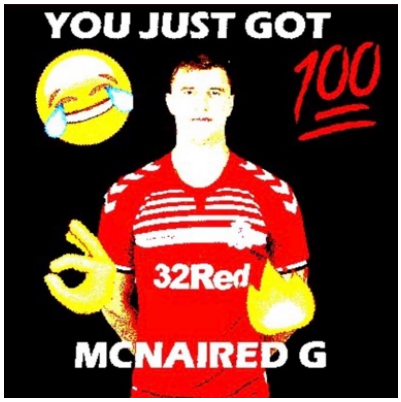
User B23 (Lines 24–26) subsequently shifts their attack strategy by returning User A23's attack back to User A23. User B23 indicates that it is User A23 who is embarrassing through the phrase "you just got" and reinforces their sarcasm through the use of an emoji combination, "💯😂👌" (i.e.,a way of saying nice meme). Since this greatly annoys User A23, User A23 (Line 27) escalates the attack further by directly mentioning that User B23 has mental problems and needs to see a doctor. In response, User B23 (Line 28) employs the same approach as in their previous turn, that is to bounce User A23's attack back to User A23. User B23 also insults User A23 for being low status, "you mong" (Line 28), and attacks User A23's gender based on User B23's own prejudice against women. User A23 (Lines 29–30) subsequently further degrades User B23 by modifying their previous description of User B23 as a keyboard warrior earlier. They state that User B23's language skills do not even qualify User B23 as a troll, "You're not even a good troll" (Line 29) and "Language skills of a 5 year old" (Line 30). In the last turn of this interaction, User B23 (Lines 31–32) uses the same approach for the third time, sending User A23's attack back to User A23. They thus criticise User A23's self-righteousness and stupidity.

Excerpt 78 (Interaction 23)

Line	User Account	Text	Level	Turn
24 25 26	User B23	[1R][PIC: https://twitter.com/CoreyMFC/status/1287072402552303616/photo/1]	12th	12
27	User A23	[1R]Hilarious. [2R:ER] [3R]Have you thought about therapy?	13th	13
28	User B23	[1R]Says the woman who doesn't believe in a virus you mong	14th	14
29 30	User A23	[1R]You're not even a good troll. [Face with tears of joy][2R:ER] [3R]Language skills of a 5 year old.	15th	15
31 32	User B23	[1R]Says the one who thinks they're cool not wearing a mask divvy mate	16th	16

Figure 26

The picture in User B23's 12th level reply



The analysis of Interaction 23 suggests that personal attacks serve to drive the development of conflictual interactions by significantly intensifying the conflict. When the recipient senses that they have been attacked, to save face, they often issue the next turn and further increase the intensity of the attack by using a stronger form of expression than the (previous) speaker. Twitter users, by taking turns being a recipient, humiliate the opponent, from attempting to restrain their opponents' freedoms of speech to attacking the opponents' characteristics (e.g., personality, intelligence level, education level) and to insulting the opponents, thus progressively intensifying the conflict.

Function 2: changing the purpose of doing the conflict

Personal attacks are prone to result in interlocutors competing for the right to the floor in the subsequent interaction: the recipient protects their right to speak whereas the speaker refuses to surrender their right. Interlocutors, with the aim of engaging in a conflictual interaction, shift from arguing about opinions to constantly struggling for discursive power. This change in purpose usually occurs in interactions elicited by the replies that contain disagreements and personal attacks because such replies are usually perceived as being used by the speaker to shift the subsequent interaction in their favour by depriving the opponent of the right to speak. For example, in Except 79, @BBCBreaking's report (Lines 01–07) on the mass layoffs and closures of Boots on the UK high street due to the pandemic evokes Interaction 24 between Users A24, B24 and C24 (see Appendix 1 for the full version).

Excerpt 79 (Interaction 24)

Line	User Account	Text	Level	Turn
01	@BBCBreaking	[1R]UK High Street pharmacy chain Boots cutting 4,000 jobs and closing 48 stores in wake of coronavirus pandemic [2R][QUOTE: https://www.bbc.com/news/business-53348519?at_campaign=64&at_custom2=twitter&at_medium=custom7&at_custom4=2764B6B6-C1D6-11EA-BC8A-8B8C96E8478F&at_custom3=%40BBCBreaking&at_custom1=%5Bpost+type%5D]	tweet	1
02				
03				
04				
05				
06				
07				

The UK companies can apply for grants from the UK government to ease the financial pressure of paying employees who stop working during the pandemic. User A24 first follows up on the @BBCBreaking's report by questioning whether Boots refunds appropriate grants, "Is @BootsUK refunding all the furlough money claimed if they're cutting 4000 jobs?" (Lines 08–09). User B24 opposes User A24, arguing that people who are dismissed are also entitled to furlough money, "People who had left their jobs and been fired were eligible to furlough"

(Lines 10–11). To oppose User B24, User A24 stresses that Boots cannot lay off staff while receiving funding from the government, “It was a Job Retention Scheme not a Furlough Fraugh Scheme” (Lines 13–14).

Excerpt 80 (Interaction 24)

Line	User Account	Text	Level	Turn
08 09	User A24	[1R]Is @BootsUK refunding all the furlough money claimed if they’re cutting 4000 jobs? @RishiSunak @hmtreasury	1st	2
10 11 12	User B24	[1R]Why would they? People who had left their jobs and been fired were eligible to furlough, it wasn’t given with the requirement that there be a job there at the end of it.	2nd	3
13 14 15 16	User A24	[1R]It was a Job Retention Scheme not a Furlough Fraugh Scheme. @hmtreasury should recover all the taxpayers money spent on furlough claims by employers who never intended to retain those jobs! #Boots #johnlewis	3rd	4

From the 4th level, User C24 replaces User A24 and takes part in the one-on-one interaction with User B24. User C24 believes that those who are made redundant will receive government fundings in other ways. They oppose User A24 by arguing that the overall government spending is not affected by whether or not Boots return the furlough money, “But the gov would have paid the same money in benefits instead” (Lines 17–18). They also point out that the amount of the grant is also reviewed by the government and is not just giving as much as the business wants, “it gave companies time to review their operations” (Line 18). User B24 disagrees, arguing that more money has to be paid for furlough than benefits, “The cost would have been less in benefits” (Line 20). In response, User C24 criticises User B24 for being too narrow-minded to understand the true purpose of furlough money, “you are looking at it on a single person basis instead of taking into account the thousands of jobs and businesses it might have actually saved” (Lines 23–25). However, User B24 (Lines 27–31) claims that there are many ways for companies to hedge their risks, implying that the return of furlough money is a matter of the company’s integrity, which enrages User C24.

Excerpt 81 (Interaction 24)

Line	User Account	Text	Level	Turn
17 18 19	User C24	[1R]But the gov would have paid the same money in benefits instead,[1WS] it gave companies time to review their operations and not make hasty quick decisions	4th	5
20 21 22	User B24	[1R]The cost would have been less in benefits. [2R]Furlough was 80% up to 2.5k. The benefit cap is £1.9k max (couple with children in London).	5th	6

(To be continued)

Line	User Account	Text	Level	Turn
23 24 25 26	User C24	[1R]But you are looking at it on a single person basis instead of taking into account the thousands of jobs and businesses it might have actually saved, most businesses would have gone bust without it, trouble is there is no real way of knowing	6th	7
27 28 29 30 31	User B24	[1R]No mate, you said they'd have paid the same in benefits which was incorrect. [2R]Most businesses would have benefited from rent and business rate holidays and could have declared temporary stops to retain employees while avoiding costs for wages. [3R]There were and are other options.	7th	8

User C24 attempts to grab the speaking right by provoking User B24's ignorance, "sounds like you are in a different country from all the businesses i know" (Lines 32–33). They also turn the interaction in their favour by questioning User B24's use of the term "Rent holidays" (Line 32). User C24 thus shifts the interaction from being about furlough money to rental holidays and also manages to bring User B24 into the topics of their own preference. Instead of surrendering the discursive power as User C24 expects, User B24 protects the right by criticising User C24 for being ignorant, "You don't know about businesses that pay rent?" (Line 34). User C24 catches User B24 in a loophole, "You said rent holidays" (Line 35), and explains that all the specific cases of rental holidays they know are failures. User C24's reply gives User B24 a critical strike and completely controls the discourse. This forces User B24 to concede and endorse User C24's explanation, "Yes, government mandated and backed rent holidays" (Line 38). However, they do not give up on clarifying their views, "other options the treasury could have considered" (Lines 39–40).

Excerpt 82 (Interaction 24)

Line	User Account	Text	Level	Turn
32 33	User C24	[1R]Rent holidays?[Rolling on the floor laughing]sounds like you are in a different country from all the businesses i know	8th	9
34	User B24	[1R]You don't know businesses which pay rent?	9th	10
35 36 37	User C24	[1R]You said rent holidays, i know lots of businesses that tried but failed to get rent holidays and without the furlough scheme would have laid off staff and possibly gone bust	10th	11
38 39 40	User B24	[1R]Yes, government mandated and backed rent holidays. I'm talking about other options the treasury could have considered.	11th	12

The analysis of Interaction 24 suggests that by combining disagreements and personal attacks, Twitter users express their own viewpoints and ensure that the rest of the interaction proceeds in their favour, thus attempting to take over the discourse. This turn design, however, also allows the recipient to easily sense the speaker's communicative intent and provokes the recipient's desire to protect their right to speak, such as User B24's response to User C24 at the 9th level in Excerpt 82. In order to avoid being deprived of their

right, the recipient tends to engage in a struggle for discursive power with the speaker in the remainder of the interaction by employing the same approach. Both sides are, therefore, involved in conflictual interactions that aim at grabbing the discursive power from each other.

Function 3: reforming the relationship between interlocutors

Personal attacks prompt the recipient to infer the (previous) speaker's personality traits, attitudes or actions based on the stereotypes of the speaker's social group and to subsequently release their own prejudice about the speaker. This change causes a consistent struggle between the interlocutors in the subsequent interaction: (1) the recipient directly or indirectly reveals their prejudice against the speaker in their own turn whereas (2) the (previous) speaker, from time to time, directly or indirectly defends themself. The speaker may state that they do not belong to the social group assumed by the recipient or by trying to break the recipient's prejudice against them, such as presenting counter evidence and mounting their own prejudice against the recipient.

A case in point is Interaction 25 that originates from @Nancy's complaint (Lines 01–04) about Apple's data cables, which is shown in Excerpt 83 (see Appendix 1 for the full version). It takes place between Users A25 and B25 regarding the quality of Apple's products. User B25 highly praises the quality, technology and design of Apple's products, to which User A25 totally objects.

Excerpt 83 (Interaction 25)

Line	User Account	Text	Level	Turn
01 02 03 04	@Nancy	[1R]As a trillion dollar company, you'd think that #Apple could figure out a way to make better cords. [2R][PIC: https://twitter.com/NotDonnaBrazile/status/1025051419907379201/photo/1]	tweet	1

Figure 27

The picture in @Nancy's tweet

As the interaction moves into the 6th level, User B25 (Lines 71–76) refutes User A25's point of view by citing their own experience of use. They object to User A25's praise of Google's Pixel system, arguing that it's easier for them to use Apple's operating system. However, the objection is taken as a personal attack by User A25 (Lines 77–81 & 111–113). In User A25's opinion, User B25 is not an ordinary Apple user as they previously believed. User A25 regards User B25 who stands up for Apple from their own usage perspective as typical of a crazy Apple "fanboy" (Line 78). Based on the stereotype of Apple fans, they form a bias and treat User B25 as a brainless, fanatical, stubborn person. After criticising User B25 for being a mindless Apple fan in the first 7th level reply (Lines 77–81), User A25 also explains why Apple's system is not as superior as Google's Pixel system in the second 7th level reply (Lines 111–113).

Excerpt 84 (Interaction 25)

Line	User Account	Text	Level	Turn
71	User B25	[1R]That's another joke. I don't spend my day customizing little widgets on my phone. I don't spend that time moving icons around in a specific pattern or rooting the phone. I use it to use apps, read email, browse the internet, use the SSH terminal. iPhone does it just fine.	6th	12
72				
73				
74				
75				
76				
77	User A25	[1R]Yea you have never used a pixel with these usual fanboy responses. I do not have to customize a thing on my day to day, but when something annoys me right I can change it. Not just accept it because that's how the operating system is made.	7th	13_1
78				
79				
80				
81				
Omit Lines 82–111				
111	User A25	[1R]by default chrome OS has an ssh terminal, and hell even Windows lets you do that with gitbash since.. Windows 7 before they adopted side loading linux	7th	13_2
112				
113				

User A25's attack on User B25 in the first 7th level reply (Lines 77–81) elicits the subsequent interaction from the 8th level to the 13th level between Users A25 and B25 about Apple fanboys. In the discussion, User B25 insists on defending themself by proving that they are not a fan of Apple. User B25 argues that they support Apple's products because of the value of the product, not because they are a fan of Apple, "How the hell am I a fanboy if I use products from all companies?" (Lines 83–84) and "someone who owns multiple products from multiple companies" (Lines 92–93). User B25 criticises User A25 for being a disrespectful and stupid person, stating that "how fragile your opinions and arguments are instead of using facts" (Lines 93–94) and "I'm talking to an idiot who doesn't realize so" (Lines 101–102). In response, User A25 extends the attack to the entire Apple fanbase by

criticising the brainlessness of Apple fans. For example, “Fanboys always go to the oh to much customization as the excuse that Apple does what they need it to do” (Lines 88–89), “my wife has an iPhone 6 regular person” (Lines 105–106) and “I have a friend loves Apple and fanboys too” (Lines 108–109). User A25 insists that User B25 is stubborn as part of the Apple fanbase. For instance, “because you are a fanboy and opinions are something you cannot change” (Lines 90–91) and “you have an opinion you are looking at this from your point of view” (Lines 95–96). User A25 also repeatedly emphasises that they wish to end the conversation, such as “I am ending this tirade” (Line 90), “I am ending this conversation” (Lines 96–97) and “I am not wasting my time with a fanboy” (Lines 97–98).

Excerpt 85 (Interaction 25)

Line	User Account	Text	Level	Turn
82	User B25	[1R]Nope, pixel’s software is not that special compared to any other android, which I have used. How the hell am I a fanboy if I use products from all companies? Nothing annoys me about iOS, what’s there to annoy me when all I do is just open apps? What do you need to change?	8th	14
83				
84				
85				
86				
87				
88	User A25	[1R]Fanboys always go to the oh to much customization as the excuse that Apple does what they need it to do. I same old song. I am ending this tirade because you are a fanboy and opinions are something you cannot change.	9th	15
89				
90				
91				
92	User B25	[1R]Calling someone who owns multiple products from multiple companies just shows how fragile your opinions and arguments are instead of using facts. Good on you.	10th	16
93				
94				
95	User A25	[1R]Dude, you have an opinion you are looking at this from your point of view, no one elses. I am ending this conversation because its fricken Twitter, and I am not wasting my time with a fanboy, who tries to displace brand love for tech innovation.	11th	17
96				
97				
98				
99				
100	User B25	[1R]I’ve been giving you facts 80% of the time. And I’m not missing out by much because I’m talking to an idiot who doesn’t realize so. Especially when you can only call me a fanboy. Same old same old “you’re a fanboy so I win” bullshit.	12th	18
101				
102				
103				
104				
105	User A25	[1R]Win what? A Twitter tirade? Listen my wife has an iPhone 6 regular person. She doesn’t want upgrade to an 8 or X because the change sucks. That’s regular person not a tech persona. I have a friend loves Apple and fanboys too, but cannot explain to me why he needs 128gb phone.	13th	19
106				
107				
108				
109				
110				

The analysis of Interaction 25 shows that replies that are perceived as humiliating the recipient can serve to change the relationship between the recipient and the speaker in the following interaction, thus progressing the interaction. Personal attacks cause the recipient to recall the stereotypes of the social group to which the speaker belongs and evoke the

recipient's personal prejudice against the speaker. This prejudice provides a basis for the recipient's view of the speaker in the ensuing conversation. In the recipient's view, the speaker is no longer just an opponent but a member of the social group that possesses many of the characteristics or personality traits that the recipient dislikes or even hates. It is as if in Interaction 25, after User A25 treats User B25 as a crazy Apple "fanboy" (Line 78) in Excerpt 84, the relationship between User A25 and User B25 shifts from being fellow Apple users to being an Apple fan and a non-Apple fan. This change in relationship also has an impact on their interactions in other streams after this. For instance, Excerpt 86 occurs after User A25 first identified User B25 as a fan of Apple (Line 77) in Excerpt 84. Having rejected User A25's criticism of Apple's battery, in order to prove the high quality of the Mac's display, User B25 (Lines 118–123) changes their approach to oppose User A25. User B25 shares their experience of using a Mac and opposes User A25 from a more objective point of view as User A25 does. In this way, User B25 counters User A25's accusations of being an Apple fanboy, by proving that they are also very objective and rational.

Excerpt 86 (Interaction 25)

Line	User Account	Text	Level	Turn
118	User B25	[1R]Almost none of this is an opinion. [2R:ER] [3R]Facts: Mac batteries can be replaced. [4R]CNET put a boulder on a Mac and submerged it in water to test its durability. [5R]Display on X is the best of all smartphones according to displaymate. [6R]X has stacked logic boards with highest density	10^{th}	22
119				
120				
121				
122				
123				

Function 4: extending the reach of interactions

The analysis shows that in interactions that have progressed to a higher level, if a reply is interpreted as a personal attack, it may suddenly receive multiple responses from Twitter users who are new to the interaction. This suggests that personal attacks located in the middle may also extend the reach of interactions by attracting more users into the interaction, just as those located at the beginning (see Function 1 in Section 6.4.1). For example, Interaction 18 (see Appendix 1 for the full version) is a one-on-one interaction between Users A18 and B18 on wearing a mask during Covid-19, as discussed in Section 5.2. After User A18 elaborates on their reasons for not wearing a mask, three new Twitter users join the interaction as can be seen in Excerpt 87, namely Users C18, D18 and E18. They all take User A18's clarification as a personal attack and criticise User A18 for being selfish, stupid and hypocritical. For instance, "someone who won't wear them in public is

selfishtwat" (Lines 37–38), "aww fuck up sean and just wear a mask you absolute nonce" (Lines 39–40) and "Shut up you fucking moron (Line 44)". This interaction thus goes from a one-on-one interaction between User A18 and their opponent, User B18, to a one-to-many interaction between User A18 and the opposing camp, including Users C19, D18 and E18.

Excerpt 87 (Interaction 18)

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
Omit Lines 04–32				
33	User A18	[1R]Cos it wasn't relevant 4 months ago and isn't now	9th	10_1
34 35 36 37 38	User C18	[1R]I bet you clapped on Thursdays for the NHS doing a great job, just to disrespect them online and potentially cause a second lockdown. You can call ppl who wear masks sheep's all you want, but I think the word for someone who won't wear them in public is selfishtwat.	10th	11
39 40 41 42 43	User D18	[1R]aww fuck up sean and just wear a mask you absolute nonce. As more and more people are now out in public and coming into contact with each other then we need to try reduce the risk of spreading it and a mask does that! Don't want to wear a mask? Then don't go out! Simple!	10th	12
44	User E18	[1R]Shut up you fucking moron	10th	13_1

However, this function is extremely rare in the corpus, mainly because: first, one-on-one interactions that reach high levels (e.g., the 9th level in Interaction 18) are rare on Twitter. Secondly, the content of such lengthy interactions is collapsed on the first page of the tweet. Twitter users have to click on the reply button multiple times to see the full conversations, which results in a low chance of new Twitter users joining such long interactions. The discussion here is only about situations like the one shown in Interaction 18, where multiple new users are inserted into a one-on-one conversation that has reached a high level. In these cases, the increase in participants is presumably due to the fact that the speaker's reply has been perceived as an attack by the recipient.

6.4.3 End

Determining the endpoint of a conflictual Twitter interaction is not an easy task. The largest scale of a conflictual interaction in the corpus is the set of all streams of message triggered by the same tweet, i.e., a CONVERSATION (see Section 4.2.1 in Chapter 4). The smallest one is a one-on-one interaction that begins with a tweet and ends at a reply that will never be replied to, i.e., a stream (see Section 4.2.1 in Chapter 4). In between, the number of adjacent streams that a conflictual interaction contains is the result of a

comprehensive consideration of various factors (e.g., research needs, participants, content of the discussion, sequence). For example, in Interaction 16 (see Appendix 1 for the full version), User A16's objection (Line 04) to @Jennifer's tweet (Lines 01–03) is futher commented on by User B16 (Lines 05–06) at the 2nd level. This leads to a one-on-one interaction between @Jennifer, Users A16 and B16.

Excerpt 88 (Interaction 16)

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A16	[1R]It's your job	1st	2
05	User B16	[1R]and it's your job as a decent human being to wear one too :)	2nd	3
06				

However, User C16 (Lines 07–09) responds to User B16 by posting two 3rd level replies. The stream that originates from @Jennifer and passes through Users A16 and B16 is split into two streams. One stream ends at User C16's first 3rd level reply (Line 07) as it is no longer being replied to. Another stream does not terminate until the 6th level.

Excerpt 89 (Interaction 16)

Line	User Account	Text	Level	Turn
07	User C16	[1R]Nope	3rd	4_1
08	User C16	[1R]If anything not wearing 1 is as healthy because people stay well clear :)	3rd	4_2
09				

User C16's second 3rd level reply (Lines 08–09) is commented on by User B16 (Line 10), eliciting a one-on-one interaction between Users B16 and C16 from the 3rd level to the 6th level. In this second reply (Lines 08–09), User C16 further reveals the reasons why they refuse to wear a mask in the first reply. This suggests that the discussion between Users B16 and C16 from the 3rd level to the 6th level is relevant to the User C16's first 3rd level reply. Interaction 16, thus, consists of two streams starting from @Jennifer's tweet: one terminates at User C16's first 3rd level reply and the other one terminates at User B16's 6th level reply.

Excerpt 90 (Interaction 16)

Line	User Account	Text	Level	Turn
10	User B16	[1R]that's... not how it works	4th	5
11	User C16	[1R]If people are happy to wear em, let em, if people are happy not wearing em, let em. not all shops are following the "guidance" and neither should they	5th	6
12				
13				

(To be continued)

Line	User Account	Text	Level	Turn
14 15	User B16	[1R]it's been advised by MEDICAL EXPERTS for a reason. they work to protect people, being a selfish twat doesn't	6th	7

As shown in Interaction 16, the end of a stream may differ from that of an interaction. The cessation of responses to a reply may not mark the end of the interaction which the reply belongs to. It is also very important to analyse the relevance of adjacent streams, such as whether the same users are engaged and whether the topic has changed. This section, thus, focuses only on replies located at the end of a stream that contain personal attacks (i.e., pure personal attacks or a combination of personal attacks and disagreements); and after these replies, the speaker and the person being attacked would also stop other one-on-one interactions prior to these replies being issued.

The analysis shows that, in most cases, such replies do terminate the interaction. As discussed in Section 6.3, not responding to the speaker's attack does not mean that the person being attacked does not sense the attack. Rather, the attacked person voluntarily ends the interaction with the speaker because of the speaker's apparent intent to attack. Therefore, after giving up responding in one interaction, the attacked person often does not engage in any further interactions with the speaker. This suggests that withdrawing has become a normal behaviour on Twitter. Users are very likely to use offensive behaviour to fight against offensive behaviour, indicating that Twitter is changing the language norm (i.e., verbal hygiene) which is generally observed in face-to-face settings. There are exceptions to this, however.

First, the attacked person sometimes launches an attack against the speaker, after giving up replying to the speaker's attack. This attack is often related to the speaker's attack, to which the attacked person does not reply. A case in point is Interaction 26 (see Appendix 1 for the full version). This interaction starts from @Brian's report on Australia bushfires (Lines 01–07) and it is between Users B26 and C26 about the cause of the Australian bushfires.

Excerpt 91 (Interaction 26)

Line	User Account	Text	Level	Turn
01	@Brian	[1R]Australia. This is the road between Canberra and the Coast I think yesterday. The idea that it is somehow not right 'to talk about #ClimateChange at this difficult time' is frankly convenient bulllshit. Click on image to see what life is like now for the #firefighters[1WS]. [2R][PIC: https://twitter.com/TwoPaddocks/status/1208207260477882369/photo/1]	tweet	1
02				
03				
04				
05				
06				
07				

Figure 28

The picture in @Brian's tweet

As the interaction moves to the 4th level, User B26 (Lines 19–22) argues that the bushfires are the result of the Australian government's inaction on climate change. They provide evidence on the severity of the problem caused by the government's failure to reduce emissions from the use of fossil fuels, "Current govt abolished it, emissions steadily risen since" (Lines 20–21). User C26 opposes User B26, arguing that population growth is the real cause of high CO2 emissions, "Rising populations are the biggest environmental issue" (Lines 26–27). In response, User B26 completely denies User C26's view, "That is utter bullshit" (Line 29), and attacks User C26 as a "coward" (Line 29) and "part of the problem" (Line 31).

Excerpt 92 (Interaction 26)

Line	User Account	Text	Level	Turn
19	User B26	[1R]Then you haven't seen enough. Our country had an ETS that visibly reduced our emissions. Current govt abolished it, emissions steadily risen since. Govt ignored business case for funding firefighting equipment for 11 mths, now not enough eqp to fight fires.	4th	5
20				
21				
22				
23				

(To be continued)

Line	User Account	Text	Level	Turn
24 25 26 27 28	User C26	[1R]We could reduce our fossil fuel emissions to zero and it wouldn't make a scrap of difference. Co2 is naturally occurring. It's everywhere and it's in everything. Rising populations are the biggest environmental issue. We are all part of the problem.	5th	6
29 30 31	User B26	[1R]That is utter bullshit. Doing nothing is the cowards way. There are plenty of solutions but the wilfully ignorant are barring the way. Don't be part of the problem.	6th	7

User C26 (Lines 32–39) arranges their objection in four 7th level replies. User C26 first defends themself by arguing that they are not ignorant, "I'm doing my part" (Line 32) and "I'm not ignorant and I do what I can" (Line 33). They then ask User B26 to provide a clear solution, "Please tell me your solutions" (Lines 33–34). They finally criticise User B26 for being unenlightened and hypocritical. For example, "Anyone can point the finger" (Line 34) and "I'm guessing none of you take unnecessary holidays, drive big cars, use air conditioners, have multiple tvs" (Lines 37–39). User B26 does not reply to any of User C26's 7th level replies and terminates the interaction via withdrawing, but it is only temporary.

Excerpt 93 (Interaction 26)

Line	User Account	Text	Level	Turn
32	User C26	[1R]I don't have any children. I'm doing my part.	7th	8_1
33 34	User C26	[1R]I'm not ignorant and I do what I can. Please tell me your solutions. Anyone can point the finger.	7th	8_2
35	User C26	[1R]Plenty of solutions[Thinking face]	7th	8_3
36 37 38 39	User C26	[1R]Nine people have liked your tweet. Not one has mentioned a solution. Typical. I'm guessing none of you take unnecessary holidays, drive big cars, use air conditioners, have multiple tvs[Face with rolling eyes]	7th	8_4

At the next level, User D26 (Lines 40–42) joins in the interaction and replies to User C26's fourth 7th level reply (Lines 36–39). They criticise User C26 for being the problem "you're part of the problem" (Line 40), emphasise that CO2 emissions can be reduced, "We have solutions" (Line 40), and propose to form a camp with User B26, "Solidarity User B26" (Lines 41–42). In response, User C26 (Line 43) provokes their opponents, including Users B26 and D26 by pressing for a specific solution once again. This seriously angers User B26. User B26 (Lines 44–48) returns to the interaction. The phrase "I debated whether or not to bother" (Line 44) suggests that User B26 does intend to end the interaction with User C26 at the 7th level. Before citing the solution, User B26 criticises User C26 for being provocative. It is evident that although User B26 voluntarily terminates the interaction at the 7th level, User B26 does feel the attack from User C26. The conflict between Users B26 and C26 has just

been put on hold rather than being resolved. Once User B26 chose to continue to interact with User C26, they still fought back against User C26's attack.

Excerpt 94 (Interaction 26)

Line	User Account	Text	Level	Turn
40	User D26	[1R]Joe, you're part of the problem. We have solutions, think big and not at the individual level. [2R:ER] [3R]Solidarity User B26[Green heart]	8th	9
41				
42				
43	User C26	[1R]Still waiting to here the solutions	9th	10
44	User B26	[1R]I debated whether or not to bother given you're not genuinely interested in the answer and are looking for a gotcha moment but here goes... [2R][QUOTE: https://www.climatecouncil.org.au/full-list-of-fire-and-emergency-chiefs-recommendations-to-federal-government/]	10th	11
45				
46				
47				
48				

Second, when the speaker's attack is not responded to by the attacked person, the speaker sometimes attacks that person again in another interaction in which that person is involved. At times, other Twitter users in the same camp as the attacked person are also attacked. For example, in Excerpt 95, @James (Lines 01–04) posts a tweet aiming to criticise the wastefulness of the rich eating gold wings. User A27 (Lines 05–07) questions @James's ulterior motives for inciting class hatred, promoting an attack from User B27 (see Excerpt 96).

Excerpt 95 (Interaction 27)

Line	User Account	Text	Level	Turn
01	@James	[1R]I can't even afford food and rich people are eating gold on theirs. Fuck capitalism and fuck the rich [2R][QUOTE: https://twitter.com/InsiderFood/status/1145157506869710848]	tweet	1
02				
03				
04				
05	User A27	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
06				
07				

At the 2nd level, User B27 (Line 08) satirises User A27's stupidity by asserting that people who can afford to eat golden chicken wings do not starve to death. The interaction between User A27 and B27 does not develop further because User A27 does not respond to User B27's attack.

Excerpt 96 (Interaction 27)

Line	User Account	Text	Level	Turn
08	User B27	[1R]WHHHHHHHHAT IN THE FRESH HELL AHAHHA	2nd	3

In response to another attack at the 2nd level in which User C27 (Line 09) satirises User A27's intelligence, User A27 defends themself by arguing that they have freedom of speech, "I

have the right of speech" (Line 10). User B27 (Lines 11–12) enters this new interaction. They attack User A27 for the second time by asserting that User A27 does not have the right to freedom of speech, "You can't even consent to having a girlfriend legally" (Line 11). This time the interaction between User A27 and B27 stops once again because User A27 does not respond to User B27's attack either.

Excerpt 97 (Interaction 27)

Line	User Account	Text	Level	Turn
09	User C27	[1R]bro you're 15	2nd	4
10	User A27	[1R]Bro I have the right of speech.	3rd	5
11	User B27	[1R]You can't even consent to having a girlfriend legally. sit down	4th	6
12				

At the 4th level, User D27 also insults User A27 (Line 10) for not having the right to free speech but the reason they give is that User A27 is a furry, "Ur a furry so no" (Line 13). In response, User E27 (Line 14) defends furries, including User A27, by criticising User D27's provocation. After two unsuccessful attacks on User A27, User B27 (Line 15) shows up for the third time. User B27 regards Users A27 and E27 as being part of the same camp and criticises User E27 as a person who changes their behaviour or opinions according to the situation, "Okay amphibi" (Line 15).

Excerpt 98 (Interaction 27)

Line	User Account	Text	Level	Turn
13	User D27	[1R]Ur a furry so no.	4th	7
14	User E27	[1R]Well that doesn't have much to do with it, but carry on.	5th	8
15	User B27	[1R]Okay amphibi	6th	9

As this interaction shows, User A27's disregard for User B27's attacks does not mitigate their conflict. User B27 feels that their attacks do not have the effect that they expect, that is to trigger a fierce counterattack from User A27. They believe that User A27 does not take their attack on User A27 seriously at all, which is a blatant provocation and defiance. So when User A27 appears once again, User B27 launches a second attack on User A27. As they are unable to sense whether or not User A27 is enraged, User B27 even extends their attack to User E27 who defends User A27.

Together, the results in this section provide important insights into the impact of personal attacks on the nature of conflictual Twitter interactions. Personal attacks located at the beginning of an interaction provide a constant topic, i.e., for or against the speaker, thus

generating relatively fixed opinion camps in the ensuing interaction. This leads to more users being involved in the interaction that becomes more aggressive. Personal attacks located in the middle escalate the interaction to a new stage of maximum intensity by greatly intensifying the conflict between interlocutors. They may provoke the recipient to protect their right to speak, thus triggering a battle for discursive power between interlocutors in the subsequent interaction. They can also change the relationship between interlocutors in the remainder of interactions by altering the recipient's understanding of the speaker's identity. The recipient views the speaker as a member of the social group that they dislike or hate rather than a user with an opposing view. At times, personal attacks may also widen the scope of interactions by attracting more Twitter users. In terms of the personal attacks at the end, they may also set a stage for re-emerging attacks between the interlocutors in a new interaction. Although, in most cases, they serve to end the interaction.

6.5 Summary and conclusion

This chapter began by quantifying the occurance and frequency of replies that are perceived as a personal attack by at least one recipient in the corpus and by investigating the reasons for such an interpretation. It found that personal attacks are commonly and widely distributed throughout the corpus across a variety of contexts. The posts that constitute a personal attack from a structural perspective in most cases serve as an indication of an escalating conflict since they clearly reveal the speaker's intent to humiliate the recipient. The recipient is, however, unable to identify personal attacks solely based on typical linguistic features such as taboo language and swear words, because their interpretation of whether posts include a personal attack is strongly constrained by the interactional context. There are a lot of personal attacks in the corpus that are employed by the speaker in an off-record manner without any overt linguistic characterisitcs that identify them as an attack. There are also many replies that do not appear to be a personal attack on the surface but which however make the recipient feel severely humiliated, including: (1) replies that give the recipient the impression that the speaker takes the moral high ground, (2) replies that make the recipient feel that their right to judge correct and incorrect is being denied, (3) replies that provoke strong annoyance on the part of the recipient and (4) replies that interfere with a key interest of the group to which the recipient belongs or that they favour.

This chapter subsequently provided an in-depth analysis of the interactional function of personal attacks, by combining both a quantitative and a qualitative approach. As Table 8 shows, the analysis found that personal attacks do serve to drive the development of conflictual Twitter interactions and seem to become an acceptable approach on Twitter to achieve certain strategic goals, involving protecting their own faces, combatting against the right to speak with the recipient and registering their objections and angers. Specifically, the analysis found that when the recipient views a post as a personal attack, they tend to attack the speaker as a response because, in their perspective, this post greatly damages their own face. In this way, personal attacks largely guarantee the occurrence of turn-taking in most cases, thus maintaining the continuation of conflictual interactions. On the other hand, personal attacks also influence the nature of conflictual Twitter interactions as a whole. First, they serve to extend the reach of interactions by increasing the number of participants. Not only personal attacks at the beginning of an interaction but also those in the middle can occasionally serve this function. Second, personal attacks function to change Twitter users' communicative intents in conflictual interactions. This is particularly evident when a reply containing both disagreements and personal attacks is perceived by the recipient as a personal attack. This reply exposes the speaker's apparent purpose to interfere with their opponents' rights to speak. It results in a consistent struggle between interlocutors to protect their own right and to seize the other's right in the ensuing interaction. Third, the most crucial role played by personal attacks in conflictual interactions is evidenced by the blatant intensification of conflict as they greatly strain the relationship between the two parties involved. They may appear anywhere in interactions and may take any form (i.e., solely disagreements or personal attacks or a combination of both). Due to the self-contained nature of the topics, personal attacks at the beginning of an interaction tend to concentrate and fix the participants in the subsequent interaction into two camps: either supporting or opposing the speaker who issued the personal attack. This avoids the fragmentation of discussions and greatly increases the aggressiveness of the subsequent interaction. In the middle, personal attacks tend to lead the recipient to attributing stereotypes about the (previous) speaker's social group to the speaker. The interpersonal conflict between the interlocutors often escalates into a more massive scope of the conflict between the communities that these interlocutors belong to and is likely to reach its climax, that is name-calling. Personal attacks located at the end may set a scene for their speakers'

next attack on either the opponent who is being attacked or this opponent's camp in a new interaction.

Table 8

A summary of findings in Chapter 6

		Previous studies	Chapter 6
Question 2: how do linguistic strategies drive the development of conflictual Twitter interactions?	Do personal attacks progress arguments?	1. Stereotype: no (see Section 5.3.4 in Chapter 5);	1. Challenging the stereotype: yes;
	Why?	1. Prevent logical objections from being raised (see Section 2.1.2 in Chapter 2);	1. Guaranteeing the occurrence of turn-taking; 2. Impacting the discursive nature of Twitter conflict: (1) Extending the reach of interactions; (2) Changing Twitter users' communicative intents; (3) Blatantly intensifying the conflict;

A further issue also emerged in this chapter. People are likely to stereotypically assume that words that are linguistically polite help to mitigate the recipient's anger. Linguistic politeness was, however, also observed in replies that are viewed as a personal attack by the recipient, such as User C2's Turn 4 (Lines 12–15) in Excerpt 59. To challenge User B2's prayer, User C2 does not target User B2's characteristics but points out a logical flaw among theists in their beliefs on the omnipotence of God, forming an indirect opposition to religion. In the meantime, they also attempt to reduce the face-damage to User B2 by using negative politeness strategies such as presenting their objection as a personal assumption, "I'm assuming that" (Line 12), and personalising their viewpoint, "to me" (Line 15). Although, literally, User B2 is not supposed to be angered and may also soften their opposition, the fact remains that User C2's Turn 4 is instead perceived by User D2 as a personal attack (Lines 16–19 & 97–99). Another case in point is User A23's Turn 6 (Lines 11–14) in Excerpt 76. Although User B23 blatantly attacks User A23 by calling them "an idiot" (Line 10), User A23 does not respond in the same manner. They quote a sentence in User B23's turn and combine it with a smiley emoji, "😂" (Line 11), and a positive assessment, "Good one" (Lines 11–12), which appears to be supportive. However, they subsequently question User B23's level of education, "Can you spell potato?" (Line 12), which hints that their prior support is fake. User A32 actually intends to insult User B23's intelligence and their use of the smiley emoji, and the positive assessment is not employed to protect User B23's positive face but more likely serves to reinforce the face-damage to User B23. Users

C2's and User A23's turns suggest that, as in face-to-face settings, people in conflictual Twitter interactions also decide whether and how to redress the damage of their objections/attacks to their opponents' faces. The two turns, more importantly, suggest that linguistic politeness seems to be used and/or interpreted differently in conflictual Twitter interactions. People appear to have reached a different consensus on verbal hygiene on Twitter. The next chapter, therefore, examines the makeup and nature of linguistic politeness and its function in conflictual Twitter interactions from the dual perspectives of how the speaker uses linguistic politeness and how the recipient interprets it.

Chapter 7 An In-depth Analysis of Linguistic Politeness

7.0 Introduction

Chapter 6 provided an in-depth analysis of personal attacks on Twitter from the perspective of the recipient. It found that personal attacks are widely and frequently distributed across different contexts throughout the corpus, and their occurrence is heavily influenced by the contexts in which the recipient interprets the (previous) speaker's turn. Contrary to common assumptions, the analysis reveals that on Twitter people are generally open to using personal attacks because they seem to regard personal attacks as a strategy that facilitates certain actions such as protecting the speaker's own face, registering their objections and anger and combatting against the right to speak with the recipient. It also found that what the recipient perceives as a personal attack are conducive to the development of conflictual Twitter interactions. These personal attacks severely threaten or damage the recipient's face, thereby encouraging further engagement and promoting the continuation of conflictual interactions. They also profoundly affect the nature of interactions that occur afterwards, including extending the reach of interactions, changing the communicative intent of Twitter users and overtly intensifying the conflict. Chapter 6 also noticed that the speaker's opposition to the recipient is not always as blatant as suggested by popular discussions of Twitter. On Twitter, people appear to use linguistic politeness to decide whether and how to redress the damage to their opponents' faces, as they do in face-to-face interactions. However, they seem to employ and perceive linguistic politeness in a partially different way because linguistic politeness was even found in replies that are perceived as a personal attack and is sometimes used to reinforce the face-hurt.

In order to obtain further insights into the use and function of linguistic politeness on Twitter, this chapter carries out an in-depth analysis of cases in which Twitter users employ linguistic politeness while rejecting their opponents' viewpoints and/or attacking the opponents. It aims to address the third question (i.e., What constitutes linguistic politeness in conflictual Twitter interactions and how does it modify the intensity of such interactions?) and to explore how technology has changed people's uses and perceptions of linguistic politeness based on Brown and Levinson's (1987/1978) facework.

Brown and Levinson (1987/1978) examined linguistic politeness largely based on the data from harmonious face-to-face interactions or interactions that seek to establish

harmonious relationships. This leads to the issue that despite a flexible explanatory model provided by their facework, linguistic politeness was only viewed as a type of linguistic strategy used by the speaker to redress potential FTAs to the recipient's face. Employing this definition in the present study is quite problematic, as mismatches between the literal and discursive functions of language use on Twitter are quite common. In order to identify a more effective way of employing the facework to conflictual Twitter interactions, Section 7.1 re-defines linguistic politeness in conflictual Twitter interactions and takes it as a modifier used by the speaker to adjust (i.e., mitigate or reinforce) the intensity of the conflict with the recipient.

Section 7.2 discusses the specific uses of linguistic politeness in the corpus. Based on Brown and Levinson's (1987/1978) taxonomy of doing FTAs, it first distinguishes the ways of doing conflict into five categories and quantifies the frequency of each method used in the corpus: (1) doing conflict bald-on-record, (2) doing conflict on-record with positive politeness, (3) doing conflict on-record with negative politeness, (4) doing conflict off-record and (5) doing conflict with pseudo-politeness (see Section 7.2.1). Following this, this section explores in more detail the makeup, discursive purpose and interactional function of each method based on a qualitative analysis of the interactions in the corpus and a questionnaire survey among Twitter users. Specifically, Section 7.2.2 investigates how Twitter users engage in a conflictual interaction overtly without redress. It finds that, as in face-to-face interactions, Twitter users hardly object to their opponents' views or attack them bald-on-record. On Twitter, however, the use of the bald-on-record approach necessarily results in severe face-damage to the recipient. Sections 7.2.3 and 7.2.4 examine how Twitter users mitigate conflict by using positive and negative politeness, respectively. Positive politeness in conflictual Twitter interactions appears to play a different role from what it does in face-to-face interactions. It tends to manage or control the conflict between interlocutors because it attempts to dissuade Twitter users from persuading their opponents to accept an opposing view. In terms of negative politeness, although the speaker attempts to use it to protect the recipient's negative face, the recipient often perceives it as inadequate or insincere or even ignores the speaker's efforts on face-saving completely. Section 7.2.5 then examines how Twitter users conduct conflict off-record. As another common approach in the corpus, the off-record approach enables Twitter users to avoid being responsible for hurting their opponents' face-needs and to preserve the opponents' faces to some extent. As evidenced

by the recipient's response, however, the off-record manner used by the speaker does not provide the recipient with an adequate sense of face-protection. Section 7.2.6 finally explores pseudo-politeness, that is, how Twitter users employ linguistic politeness to aggravate conflict. It demonstrates that the speaker's pseudo-politeness, in most cases, irritates the recipient, triggering the recipient's attacks. It also recognises that although pseudo-politeness is also used in face-to-face arguments, it is indeed more evident in conflictual Twitter interactions because it maintains the uninhibited nature of CMC.

Section 7.3 compares linguistic politeness used in conflictual Twitter interactions with Brown and Levinson's (1987/1978) notion in more depth and attributes the formation of the stereotype that Twitter lacks courtesy to three possible reasons. First, based on their experiences of communicating with others in face-to-face contexts, people, including Brown and Levinson (1987/1978), incorrectly interpret being friendly and kind as the primary interactional goal and mistakenly treat linguistic politeness as a strategy for accomplishing this goal. However, the examination of the use of linguistic politeness in conflictual Twitter interactions in which face-damage is quite common suggests that the strategic goals and functions of linguistic politeness on Twitter only partially overlap with those outlined by Brown and Levinson (1987/1978). Second, although in most cases, positive politeness functions to manage or even resolve personal conflict between interlocutors, it is rarely applied in conflictual Twitter interactions because it requires the speaker to refrain from asserting both their dominance and fighting to be heard. Other politeness modifiers that are used to redress FTAs, including negative politeness and 'off-record' politeness, fail to make the recipient feel that their face has been adequately preserved, as shown in Section 7.2. Third, Twitter algorithms provide an opportunity to demonstrate that a very low number of bald-on-record oppositions and pseudo-politeness approaches tend to have a profound effect on people's perception of Twitter. This creates the stereotype, raises the threshold for courtesy on Twitter and results in people's linguistic behaviour being much more difficult to interpret as polite, thereby reinforcing the stereotype that Twitter is rude.

7.1 Linguistic politeness in conflictual Twitter interactions

Brown and Levinson (1987/1978) acknowledged Grice's cooperative principles (CPs) (1975) and argued that being polite is a powerful motivator for breaking the CPs. The use of linguistic politeness allows individuals to produce utterances that are no longer informative,

valid, relevant or explicit. Although they did not provide an explicit definition of linguistic politeness, they expounded it based on a set of assumptions about conversational logic. They interpreted linguistic politeness within a model called facework. They regarded face as individuals' basic needs and assumed that everyone has positive face (i.e., the want to be recognised) and negative face (i.e., the want to not be disturbed). They also presumed that every interaction bears the possibility of carrying FTAs and insisted that the FTAs undermine interactions and participants' relationships. Linguistic politeness is thus viewed as a type of linguistic strategy used by the speaker to redress possible FTAs against the recipient's face.

The authority of politeness in Brown and Levinson's (1987/1978) work is the speaker, which is quite crucial in the present study because speakers and recipients assume different responsibilities in conflictual Twitter interactions: the speaker needs to employ linguistic politeness and the recipient needs to interpret the speaker's strategies and to generate their own discursive goals based on their interpretation. The recipient may not feel that their face has been adequately protected by the speaker's linguistic politeness (Fraser, 1990); even if the speaker's actions contrast with a norm (e.g., publicly humiliating the recipient) and the speaker does not redress their FTAs, the recipient may still not evaluate the speaker's words as impolite (Fraser, 1990). Therefore, strictly differentiating the speaker's use of linguistic politeness from the recipient's interpretation of that usage helps to overcome the drawback of Brown and Levinson's (1987/1978) work (see Section 3.4.3 in Chapter 3). That is, they maintained that there is a one-on-one relationship between linguistic strategies and their communicative intent and impact. For example, in Excerpt 99, User C2's (Lines 12–15) objection to User B2's (Lines 10–11) theistic opinion, as stated by Brown and Levinson (1987/1978), puts some pressure on User B2 to abandon their theistic view and thus damages User B2's negative face. User C2 also employs negative politeness to redress this potential damage in Turn 4. They use hedges such as "I'm assuming that" (Line 12) and "to me" (Line 15) to personalise their opposition. They also change the reference to User B2 with the indefinite pronouns, "everyone" (Line 12) and "no-one" (Line 13), to phrase their objection as a general statement and to provide explanation of their view. Responding to the attack, User B2 (Lines 16–19), however, overtly attacks User C2 by threatening that User C2's "day of reckoning will come" (Line 16). They continue to satirise User C2 as a hypocritical atheist and end their turn with another threat. They threaten User C2, saying that User C2 will give up their position one day, "But that would never be you, right??? You'll have to wait

and see…" (Lines 18–19). User B2's reaction suggests that User B2 does not treat User C2's negative politeness as face-protection.

Excerpt 99 (Interaction 2)

Line	User Account	Text	Level	Turn
10	User B2	[1R]And the animals and wildlife too, Lord, created by you.[Folded hands][Folded hands][Folded hands]	2nd	3
11				
12	User C2	[1R]I'm assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn't sound like a very compassionate deity to me.	3rd	4
13				
14				
15				
16	User B2	[1R]Relax Atheist, your day of reckoning will come... everyone's does. But it's funny how, when one's life or a loved one's life is in danger people turn to asking their God to save them. But that would never be you, right??? You'll have to wait and see...	4th	5_1
17				
18				
19				

Brown and Levinson (1987/1978) argued that linguistic politeness refers to what the speaker communicates to the recipient, rather than an inherent feature of the speaker's utterance. This perspective of the use of linguistic politeness approaches (as being dependent on the interactional context rather than having fixed meanings/discursive goals) is highly relevant to this study because people's uses of linguistic politeness are not fixed in conflictual Twitter interactions. Twitter users are also likely to employ linguistic politeness to reinforce the damage to the recipient's face. For instance, apologising to express sympathy is a conventional form of politeness. Brown and Levinson (1987/1978) also refer to it as positive politeness as it serves to claim common ground with the recipient. In Excerpt 100, @washingtonpost (Lines 01–09), a famous American news media, reports the death of Qasem Soleimani. They describe Qasem Soleimani, who is regarded as a terrorist by the American government, as "Iran's most revered military leader" (Lines 01–02). User A28 appears to treat the death of Qasem Soleimani as the loss of @washingtonpost, expressing their sympathy for @washingtonpost, "We're so sorry for your loss" (Line 10). They, however, then turn to criticise @washingtonpost's incorrect political stance, "How was he related to you again?" (Lines 10–11). This suggests that User A28 actually argues that @washingtonpost is not supposed to describe Soleimani with the adjective "revered" (Line 02) because it is not in line with the USA national interest. Thus, User A28's apology here is used to intensify the sarcasm to @washingtonpost, aiming to satirise @washingtonpost as a traitor to the USA, rather than protecting @washingtonpost's face.

Excerpt 100 (Interaction 28)

Line	User Account	Text	Level	Turn
01	@washingtonpost	[1R]Breaking news: Airstrike at Baghdad airport kills Iran's most revered military leader, Qasem Soleimani, Iraqi state television reports [2R][QUOTE: https://www.washingtonpost.com/world/national-security/defense-secretary-says-iran-and-its-proxies-may-be-planning-fresh-attacks-on-us-personnel-in-iraq/2020/01/02/53b63f00-2d89-11ea-bcb3-ac6482c4a92f_story.html?utm_campaign=wp_main&utm_medium=social&utm_source=twitter]	tweet	1
02				
03				
04				
05				
06				
07				
08				
09				
10	User A28	[1R]We're so sorry for your loss, WaPo. How was he related to you again?	1st	2
11				

Brown and Levinson (1987/1978, p. 7) also presupposed "the nature of communication as a special kind of intent designed to be recognised by the recipient". They maintained that "to mean something by an utterance is to intend to produce some effect in an audience by means of the recognition of this intent" (Pfister, 2010, p. 1273). In 1987, they further clarified that when referring to the speaker's intent, they mean the speaker's communicative intent rather than their psychological one. For example, to compliment the recipient on their new haircut, the speaker may show that they have noticed the recipient's change and thus support their positive face. However, this does not necessarily mean that the speaker sincerely likes the recipient's new haircut. The speaker's communicative intent refers to what they aim to highlight in an utterance, which may, however, differ from what they really think (Brown & Levinson, 1987/1978). The use of linguistic politeness merely projects a kind or friendly exterior but is not necessarily equal to their inner feelings. The recipient has to interpret the speaker's intent based on the discursive context. If the contextual features cohere with the default meaning of politeness approaches, the speaker is perceived to be genuine; otherwise, the speaker is regarded as unauthentic. For instance, in Excerpt 101, using an indicator of disagreement, "Well" (Line 02), and showing two contradictory terms, "yes and no" (Line 02), User B indicates that "[they] cannot be telling the truth" (Brown & Levinson, 1987/1978, p. 221). In order to prevent User A's face from being hurt by their direct and accurate expression of feelings, they do the FTA in an off-record manner.

Excerpt 101 (Brown & Levinson, 1987/1978, p. 221)

Line	Speaker	Text
01	User A	Are you upset about that?
02	User B	Well, yes and no.

Taking a similar approach, User A28 in Excerpt 100 invites a conversational implicature by contradicting their sympathy, “We’re so sorry for your loss” (Line 10), with their criticism, “How was he related to you again?” (Lines 10–11). Although they may have @washingtonpost’s face in mind, preserving @washingtonpost’s face is unlikely to be their key motivation for phrasing their turn in such a manner because newspapers are generally not associated with face-wants, and people are unlikely to care about such institutions’ face-wants. User A28’s contradiction thus serves to intensify the damage to @washingtonpost’s face and is thus best interpreted as a kind of ‘fake’ politeness.

Linguistic politeness in the corpus, however, cannot serve to reveal Twitter users’ primary communicative intent in the same way as conflictual linguistic strategies. This creates the greatest inconsistency with Brown and Levinson’s (1987/1978) study that regards the nature of linguistic politeness as a type of linguistic strategy. For example, Brown and Levinson (1987/1978, p. 104) argued that the sentence “what a fantastic garden you have” is more polite than the one “you have a fantastic garden” because the exclamatory sentence shows the speaker’s strong approval of the garden that the recipient is proud of. The first sentence is more effective in establishing a rapport with the recipient than the latter one. However, if the speaker primarily aims to exchange objective information, both sentences are suitable for expressing the speaker’s positive assessment of the recipient’s garden. In this case, being polite is not as critical as clearly conveying a message. This situation is particularly evident in conflictual Twitter interactions. For instance, in Excerpt 102, User A29 challenges the credibility of @BBCBreaking’s report of the impact of fire on the structure of Notre-Dame, arguing that “a structural engineers’ report” (Lines 03–04) is more reliable. They also personalise their objection through the use of the hedge “I think I’d” (Line 03) to preserve @BBCBreaking’s negative face. Whether or not this hedge is added, User A29’s Turn 2 functions as a denial of @BBCBreaking’s point of view because the conventional expectation is that an agreement involves boosting or asserting a similar degree of importance; but here User A29 reduces the importance by referring to “structural engineer” (Line 03), that is people who are more specialised in structural issues than “Paris fire officials” (Line 02). This suggests that in conflictual interactions, linguistic politeness serves as a modifier to adjust the intensity of conflict, rather than a strategy that can alter the speaker’s primary communicative intent.

Excerpt 102 (Interaction 29)

Line	User Account	Text	Level	Turn
01 02	@BBCBreaking	[1R]Structure of Notre-Dame saved from total destruction, Paris fire official says https://bbc.in/2UAMlVn	tweet	1
03 04	User A29	[1R]I think I'd be inclined to wait for a structural engineer's report to be fair.	1st	2

Considering all this evidence, this study regards linguistic politeness in conflictual Twitter interactions as a modifier that serves to adjust the intensity of conflict between interlocutors. As can be seen in Table 9, The speaker may use it to mitigate or reinforce the conflict. However, the recipient may not be able to recognise the speaker's communicative intent to preserve their face or may not perceive that their face has been adequately protected and as a result, enters the conflict with the speaker. The next section examines linguistic politeness in conflictual Twitter interactions in greater detail.

Table 9

A comparison of parameters of linguistic politeness used in face-to-face interactions and in conflictual Twitter interactions

	In face-to-face interactions	In conflictual Twitter interactions
Who uses linguistic politeness?	The speaker;	The speaker;
What does linguistic politeness project?	The speaker's communicative intent;	The speaker's communicative intent;
Why is linguistic politeness employed?	To redress FTAs;	To redress and/or to intensify FTAs;
What is the role played by linguistic politeness?	Linguistic strategy;	Modifier;
Who is the addressee of linguistic politeness?	The recipient;	The recipient;
How is linguistic politeness interpreted?	Based on discursive context;	Based on discursive context.

7.2 Specific use of linguistic politeness

This section explores in detail the use of linguistic politeness in the corpus. Section 7.2.1 developed a model for linguistic politeness in conflictual Twitter interactions based on Brown and Levinson's (1987/1978) taxonomy and quantifies the frequency of different types of linguistic politeness in the corpus. This section discusses each type in more detail, respectively, including their makeup, natures and roles in the interaction.

7.2.1 Classification and quantification

Brown and Levinson (1987/1978) proposed three types of linguistic politeness and a possibility of being polite in different situations where the speaker may damage the

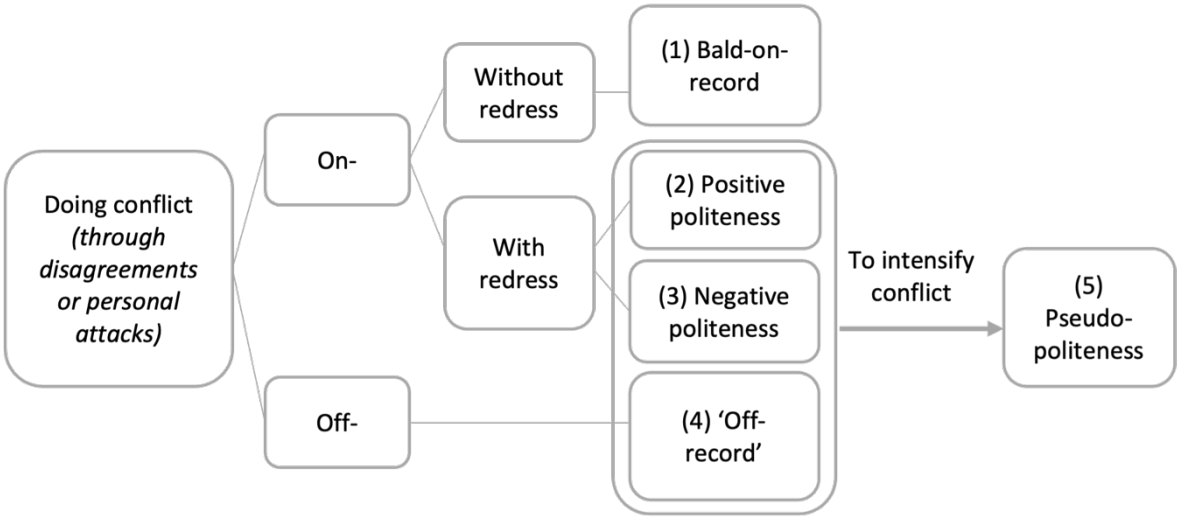
recipient's face. They first distinguished different ways of doing FTAs: the speaker gives up doing FTAs or does FTAs in an off-record or an on-record manner. In cases where the speaker conducts FTAs indirectly, ambiguously and uncertainly, they approach the FTAs in an off-record manner. If the speaker makes FTAs directly, clearly and concisely, they go on-record to do the FTAs. The speaker may also utter on-record without redress (i.e., bald-on-record manner) or proactively reduce the potential damage to the recipient's face with redress (i.e., positive politeness, negative politeness). Based on these situations, Brown and Levinson (1987/1978, p. 2) identified "three main strategies of politeness". When doing FTAs in an on-record manner, the speaker is often perceived as being polite if they mitigate the damage to the recipient's negative or positive face. The redress is, accordingly, divided into positive politeness and negative politeness. Another type is the so-called "off-record (politeness)" (Brown & Levinson, 1987/1978, p. 2), but they argued that there are limitations to this type. Politeness is only used when the speaker does the FTA off-record with the purpose of preserving the recipient's face. In this case, the speaker wants to perform an FTA but avoids the responsibility for it and the recipient is unable to "attribute only one clear communicative intent to the speaker's act" (Brown & Levinson, 1987/1978, p. 211). As for doing FTAs bald-on-record, although it generally causes the most serious damage to the recipient's face, there are exceptions. For example, both the speaker and the recipient clearly understand the intent of communication; the speaker prefers the efficiency over the recipient's face; the speaker determines that the danger towards the recipient's face is tiny, and the speaker is superior in power to the recipient. Brown and Levinson (1987/1978) claimed that in such cases the requirements for what counts as polite are comparatively low. Even if the speaker goes on-record to utter the statement without any redress, the recipient may not feel offended.

Based on Brown and Levinson's (1987/1978) taxonomy, this study further examined the conflictual linguistic strategies (i.e., disagreement and personal attacks) in more depth and identified five ways of doing conflict on Twitter and four types of linguistic politeness, which can be found in Figure 29. The analysis shows that Twitter users may go bald-on-record to object to their opponents' opinions or attack their opponents. Twitter users may also employ three types of linguistic politeness to mitigate the conflict with their opponents: (1) if the speaker protects the recipient's positive face while clearly, concisely and directly rejecting the recipient's opinion or attacking the recipient, they mitigate the conflict with

opposing the recipient, they mitigate the conflict through negative politeness and (3) if the speaker objects to the recipient's opinion or attacks the recipient indirectly, vaguely and indefinitely, they mitigate the conflict in an off-record manner. The corpus also suggests that the above-mentioned linguistic politeness, when certain conditions are met, may also serve to reinforce the conflict in interactions. This type of 'fake' linguistic politeness is referred to as *pseudo-politeness* in this study.

Figure 29

Possible ways of doing conflictual Twitter interactions



This study tried to quantify each type of linguistic politeness in the corpus by taking a similar approach to that used in Chapter 6. A large-scale quantification of linguistic politeness failed, however, because of its ambiguity and complexity in interactions. First, linguistic politeness, in most cases, appears not to be characterised by fixed linguistic features, and the features that seem to fulfil this role usually function as different types of linguistic politeness depending on context; in some cases, identification based on discursive context is also problematic. A case point is rhetorical questions. They can either be used as negative politeness or 'off-record' politeness in conflictual Twitter interactions. For example, in Excerpt 103, User A30 personalises their opinion, "I have difficulty in believing" (Line 04), when questioning the cause of the Notre Dame fire as provided by @BBCBreaking. Following this, an explanation is provided, "Considering the damage and the speed with which the fire developed" (Lines 03–04), which is designed to protect @BBCBreaking's negative face. In response, User B30 (Lines 05–07) challenges User A30, arguing that the Notre Dame fire is a total accident. They raise two rhetorical questions successively, "Really?" (Line 05) and "It's that hard to believe a really old mostly wooden interior with draperies, tapestries and wood

pews could go up in flames that quickly?" (Lines 05–07). According to Brown and Levinson (1987/1978), these questions are the manifestations of different types of linguistic politeness. The first one serves as 'off-record' politeness because it is more indirect than a blatant denial and allows User A30 to re-check their opinion. The latter one may serve to protect User A30's negative face because it provides reasons for why Notre-Dame fire was not an accident and shows that User B30 respects User A30's role as a speaker. On the other side, since the reason it provides is quite specific and clear, the second rhetorical question may also serve as an irony to criticise User A30 for suppressing evidence in an off-record manner.

Excerpt 103 (Interaction 30)

Line	User Account	Text	Level	Turn
01 02	@BBCBreaking	[1R]Structure of Notre-Dame saved from total destruction, Paris fire official says https://bbc.in/2UAMlVn	tweet	1
03 04	User A30	[1R]Considering the damage and the speed with which the fire developed, I have difficulty in believing this is purely accidental	1st	2
05 06 07	User B30	[1R]Really? It's that hard to believe a really old mostly wooden interior with draperies, tapestries and wood pews could go up in flames that quickly?	2nd	3

Second, different types of conflictual linguistic strategies in the corpus are parcelled into different sentences. These sentences, however, do not always neatly fit any of the five methods: (1) doing conflict bald-on-record, (2) doing conflict on-record with positive politeness, (3) doing conflict on-record with negative politeness, (4) doing conflict off-record and (5) doing conflict with pseudo-politeness. The corpus shows that people sometimes layer one type of linguistic politeness on top of another to further weaken or intensify face-damage. A case in point is Excerpt 104, where User A31 attacks @Matthew by going off-record and using negative politeness simultaneously. @Matthew, an environmental activist, reports on the climate strike in Berlin in the tweet. In response, User A31 overtly satirises the participants of the climate strike as "Zombies" (Line 04). Since @Matthew identifies himself as a supporter of these participants, User A31's sarcasm constitutes an off-record personal attack on @Matthew. They also further redress the damage to @Matthew's negative face by weakening the assertion of their attack, "Looks like" (Line 04), and personalising their view, "to me" (Line 04).

Excerpt 104 (Interaction 31)

Line	User Account	Text	Level	Turn
01	@Matthew	[1R]An estimated 100,000 people in Berlin on #ClimateStrike. Today is massive [2R][VIDEO: https://twitter.com/i/status/1175026670593597441]	tweet	1
02				
03				
04	User A31	[1R]Looks like a bunch of Zombies to me.	1st	2

This study, therefore, only measured the frequency of linguistic politeness that is typical of conflictual Twitter interactions and is more appropriate for identifying the changes in the intensity of conflict. Overall, only around 3% (342) of the posts in the corpus are bald-on-record, which verifies Brown and Levinson's (1987/1978, p. 95) observation that "the majority of natural conversations do not proceed in such a [bald-on-record] fashion at all". They also maintained that the bald-on-record manner may also be considered polite by the recipient if both the speaker and the recipient are clear about the purpose of their communication, if the speaker prefers efficiency over the recipient's face, if the speaker judges that the danger to the recipient's face is tiny and if the speaker is more powerful than the recipient. However, these cases were not identified in the corpus. That is, bald-on-record conflict in conflictual Twitter interactions invariably poses serious damage to the recipient's face. All of the remaining posts in the corpus (97%; 10884) contain linguistic politeness which challenges the stereotype that there is no politeness on Twitter. This linguistic politeness includes the types that are used to mitigate conflict, namely positive politeness, negative politeness and 'off-record' politeness, pseudo-politeness, which is used to intensify conflict. There are fewer than 0.5% (56) of the posts in the corpus that contain positive politeness, but nearly all of them are useful for mitigating or managing the personal conflict between the speaker and the recipient. They are, however, less effective in de-escalating the entire interaction. There are also approximately 0.7% (70) of the posts in the corpus that include pseudo-politeness. In spite of its small percentage, pseudo-politeness does serve as indicative of Twitter and is invariably taken as a personal attack by the recipient. As for negative politeness and 'off-record' politeness, almost all replies that used linguistic politeness in the corpus consist of one of them or contain both of them, such as User A38's Turn 2 in Excerpt 113 which will be discussed in Section 7.2.4. In the majority of cases, the recipient does not feel that their face is being protected due to the use of negative politeness and/or 'off-record' politeness.

The remainder of this section explores in detail the bald-on-record conflict strategy and each type of linguistic politeness modifier. The discussion is mainly based on

CONVERSATION 17 but also draws on other CONVERSATIONs to illustrate the different types. CONVERSATION 17 is triggered by @Jessica's tweet (see Excerpt 105), where she invites women to participate in a discussion about, "What would you do if all men had a 9pm curfew?" (Lines 01–02). In the second sentence, she changes the addressee from women to men and emphasises that men are not allowed to participate in this conversation. This constraint is further stressed at the end with hashtags. @Jessica reaffirms that this discussion is designed for women and the topic is about male privilege. In her tweet, @Jessica addresses men and women separately. To elicit women's views, she preposes the question for women with "Ladies" (Line 01), which is a formal and courteous synonym for women, referring to "[women] of superior social position and of refinement and gentle manners" ("Merriam-Webster—Lady", 2022).[27] However, men's views are requested by preposing "dudes" (Line 02) to the request: "Dudes: Read the replies and pay attention" (Lines 02–03). Dude is a vocative often used by men, especially men at a young age, in colloquial settings, to refer to each other. Referring to men in an informal manner is likely to create a more relaxed atmosphere, thus redressing the damage to men's negative face that is caused by the subsequent overly formal requests. When comparing the sentences addressed to women and men, it is clear that men's face is being more seriously damaged because women are asked to participate by giving their views while men are asked to listen, i.e., told to do nothing interactive. @Jessica also further redresses the face hurt to women because she blurs her role as the authority of the request by saying "a question for you" (Line 01). The interactions in CONVERSATION 17 involve almost all types of linguistic politeness in the corpus. The opinion camps in all interactions are also very clear: (1) Twitter users representing women versus (2) Twitter users representing men. This further makes the analysis of politeness less subject to the speaker's language, culture, age and other demographic characteristics that commonly play a role in face-to-face settings but are difficult to capture on Twitter.

Excerpt 105

Line	User Account	Text	Level	Turn
01 02 03 04	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1

[27] Accessed, 22nd February 2022: https://www.merriam-webster.com/dictionary/lady

7.2.2 Doing conflict bald-on-record

According to Brown and Levinson (1987/1978), doing conflict bald-on-record refers to Twitter users rejecting their opponents' opinions or attacking them in the most straightforward, unambiguous and concise manner. A case in point is User B23's objection to User A32's Turn 2 in Excerpt 106. User A32 (Lines 05–06) criticises @Jessica's tweet of deliberately provoking men and phrases their criticism as a rhetorical question. They thus protect @Jessica's face by avoiding publicly denying @Jessica's view. At the same time, they also use the auxiliary verb to "do" (Line 05) and the tag question, "right?" (Line 06), to emphasise their annoyance. Following this, User A32 overtly attacks @Jessica for being "condescending and shallow" (Line 07). However, they also disassociate themself from the attack through the verb phrase, "come off" (Line 06), indicating that the attack is not intentional but merely a description of the effect of @Jessica's post. The attack is thus further intensified by this ambivalent expression. Finally, User A32 (Lines 07–09) requests that @Jessica revises the tweet from being about the men's curfew to referring to woman's self-protection. In order to protect @Jessica's negative face, they word the request as a suggestion by using "how about" (Line 07) as this avoids sounding like they are ordering @Jessica. User B32 overtly rejects User A32's opposition, arguing that "you clearly missed the point" (Line 10). If compared to sentences such as "sorry, I might not have made myself clear" and "sorry, I think you misunderstood what I was trying to say", this statement appears to have an overtone that criticises User A32 for being a bit dumb. It is, however, not employed as an attack by User B32 because they just attribute all of the faults for the misinterpretation to User A32 rather than deliberately humiliating or insulting User A32.

Excerpt 106 (Interaction 32)

Line	User Account	Text	Level	Turn
05	User A32	[1R]You do realize the vast majority of dudes don't bother or harrass women right? Your post comes off as incredibly condescending and shallow. How about instead of generalizing you start encouraging women and tell them how to protect themselves if they would like the option?	1st	2
06				
07				
08				
09				
10	User B32	[1R]Dylan... You clearly missed the point.	2nd	3

Users A33's and C33's Turns 3 and 4 in Excerpt 107 provide a typical example of bald-on-record attacks. User A33 wishes to take "Long, calm, solo nighttime walks" (Line 05) during the men's curfew, which is opposed by User B33. User B33 goes off-record to attack User A33. They imply that women are at the origin of gender violence by telling User A33

(and all women) to dress 'properly', "you can have a long solo walk …put something on other then 90% nude" (Lines 06–07). However, this attack constitutes a serious insult and attack on User A33 because User B33 implies that User A33 and women in general are exhibitionists and thus cause gendered violence. Taking this approach, User B33 also justifies their attack and shirks responsibility for it. At the end, User B33 passes the buck again, saying "like how dear [dare] you blame men for this" (Line 07), which further provokes User A33. User A33 responds to User B33 with a bald-on-record attack. They overtly insult User B33, "an idiot" (Line 08), and publicly threaten them by arguing "I'm blocking you" (Line 08). At the next level, User C33 enters the interaction and overtly attacks User A33 by also declaring "I'm blocking you" (Line 09). They also openly support User B33's attack on User A33 by asserting that "I like this game" (Line 09), which further attacks User A33 in an off-record manner.

Excerpt 107 (Interaction 33)

Line	User Account	Text	Level	Turn
05	User A33	[1R]Long, calm, solo nighttime walks. That would be amazing.	1st	1
06 07	User B33	[1R]you can have a long solo walk ...put something on other then 90% nude -.- like how dear you blame men for this -.-	2nd	2
08	User A33	[1R]You're an idiot and I'm blocking you.	3rd	3
09	User C33	[1R]and i'm blocking you. I like this game.	4th	4

The quantitative assessment suggests that Twitter users rarely go bald-on-record to do conflict as only approximately 3% (342) of the posts in the corpus are bald-on-record objections or/and attacks, and only about half of these posts strictly coincide with Brown and Levinson's (1987/1978) definition of bald-on-record. They argued that bald-on-record utterances completely conform to every piece of Grice's CPs. Grice (1975, p. 26) assumed that the speaker needs to "make [their] conversational contribution, such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which [they] are engaged". However, in both face-to-face and Twitter interactions, there are few utterances that meet all four maxims simultaneously (i.e., the maxims of quantity, quality, relevance and manner). The CPs propose an ideal situation and/or a context-free assumption that serves as a context for interpreting people's conversations. However, it omits the fact that communication takes place in particular contexts and that there are always different options, each of which carries meaning. In conflictual Twitter interactions, the bald-on-record manner allows the speaker to unambiguously and accurately convey the message about whether they are rejecting the recipient's view or attacking the recipient.

The recipient can also easily and clearly interpret this message. However, it is clear from the corpus that what the speaker wants to communicate is often more than that. Relational content such as the speaker's mood, virtual identities and relationships to the surroundings is likewise an important part of the interaction. Interactions are therefore rarely as the CPs set out. For example, although the analysis found that User B32's Turn 3 in Excerpt 106 opposes User A32 bald-on-record, it actually still violates the CPs because the booster "clearly" (Line 10) violates the maxim of quality. This adverb serves to stress User B32's certainty of their opposition. Relational content is also appropriate for clarifying the speaker's point of view. For example, in Excerpt 105, @Jessica repeatedly highlights that the discussion is about male privilege among feminists. She asks directly for women's participation, reminds men not to participate and uses hashtags to highlight the key words once more. Although @Jessica's turn violates the maxim of quantity, it is this constant emphasis on the key point that leads all interactions in CONVERSATION 17 to revolve around the topic of gender privilege and thus serves to keep the interaction coherent and on-topic.

7.2.3 Doing conflict on-record with positive politeness

In harmonious interactions, doing FTAs on-record with redressive actions means that the speaker does FTAs publicly while giving face to the recipient (Brown & Levinson, 1987/1978). Through redress, the speaker attempts to offset the potential damage that their FTAs may cause to the recipient's face. They aim to demonstrate clearly that they do not intend or desire to threaten the recipient's face and to imply that they generally acknowledge the recipient's face and wish that their face is respected (Brown & Levinson, 1987/1978). One of the two ways of performing redress is to preserve the recipient's positive face; this is discussed here. The other one is to protect the recipient's negative face, which will be discussed in Section 7.2.4.

By using positive politeness, the speaker implies that they view the recipient's face as desirable. Disagreement studies based on face-to-face interactions (e.g., Muntigl & Turnbull, 1998; Rees-Miller, 2000; Kakavá, 2002; Ben-Menachem & Livnat, 2018) have established that speakers commonly employ positive politeness to reduce the damage caused by disagreements to the recipient's face. There are around 4% (424) of the posts in the corpus that share the same or similar linguistic features to those identified as positive politeness by Brown and Levinson (1987/1978). An in-depth analysis, however, suggests that 87% (368) of

these posts fail to preserve the recipient's positive face in both personal attacks and disagreements. The interaction fragment in Excerpt 108 is a case in point.

Excerpt 108 (Interaction 34)

Line	User Account	Text	Level	Turn
229 230 231 232	User E34	[1R]No one is fine with curfew. You aren't, I'm not. The problem is that you were asked to think and listen to women. We would do mostly simple stuff that you take for granted. That's a lesson, I think. Just listen.	20th	64
233 234 235 236 237	User D34	[1R]I understand that women feel threatened and scared, I get that. But I think that is terribly wrong to start that debate with - lets curfew all men. That message just makes feminists seem like man haters and doesn't even allow us to debate the issue that is underneath all of this.	21st	65
238 239 240 241	User C34	[1R]ok honestly I get what youre saying but I think some really did take it too literal. I[Face with tears of joy]agree compltly with what youre saying but I also agree with what I am getting at..i have no idea if any of this makes sense but yh	22nd	66

User E34 first provides the previous speaker (i.e., User D34) with common ground on curfew, "No one is fine with curfew. You aren't, I'm not" (Line 229), which is treated as positive politeness by Brown and Levinson (1987/1978). User E34 (Lines 230–231) then clarifies that the purpose of this conversation is not to establish a men's curfew but to make men aware of the hardships they impose on women and to encourage them to show respect for women. These explanations protect User D34's negative face. At the end, User E34 emphasises that men should "listen" (Line 232) to what is being discussed. Before raising an objection, User D34 expresses their concern for women's needs through a sentence beginning with the phrase "I understand that" (Line 233). This is used to assert a partnership or common ground with User E34, which is further emphasised by the statement, "I get that" (Lines 233–234). These explanations of their position also suggest that User D34 is attempting to avoid a disagreement, which is believed by Brown and Levinson (1987/1978) as protecting User E34's positive face. After a conjunction, "But" (Line 234), User D34 directly denies User E34's opinion on the men's curfew, arguing that using the assumption of men's curfew is a form of discrimination against men by feminists. The response of User C34 indicates their notice of User D34's comment, "ok" (Line 238), and emphasises their concerns of User D34's view, "honestly I get what youre saying" (Line 238), which is taken as positive politeness by Brown and Levinson (1987/1978). They then criticise User D34 for being too doctrinaire, "take it too literal" (Line 239). User C34 also sticks to their own view, "but I also agree with what I am getting at" (Line 240) while explicitly acknowledging User D34's opinion "i[Face with tears of joy]agree compltly with what youre saying" (Lines 239–

240). According to Brown and Levinson (1987/1978), the adverb “compltly” (Line 239), serves as positive politeness because it boosts User C34’s approval of User E34. A similar role is played by the emoji, “😂” (Line 239), which simulates the habit of breaking up a complete sentence with laughter in spoken language, thus easing awkwardness or showing kindness. User C34 ends the turn by stating that the above statement is a personal position and gives User D34 the option to disagree, “..i have no idea if any of this makes sense but yh” (Lines 240–241). This approach, together with the use of the hedge “I think”, recurring in Excerpt 108 (e.g., Lines 231–232, 234, 238), is typical of negative politeness, which will be discussed in the next section.

According to Brown and Levinson (1987/1978), there are a total of four utterances in Excerpt 108 that employ positive politeness, which are:

(1) “No one is fine with curfew. You aren’t, I’m not. The problem is…” (Line 229);

(2) “I understand that women feel threatened and scared, I get that. But..” (Line 233);

(3) “ok honestly I get what youre saying but…” (Line 238);

(4) “i[Face with tears of joy]agree compltly with what youre saying but… I have no idea if any of this makes sense but…” (Lines 239–241).

The speakers in (1) and (4) claim common ground by arguing that they share similar opinions with the recipient. The speakers in (2) and (3) state the partnership with the recipient by presenting their concerns for the recipient’s opinion. If the four sentences were derived from harmonious interactions, they are undoubtedly constructed with positive politeness. However, the problem is whether the four sentences are still used by the speaker to preserve the recipient’s positive face in conflictual interactions. The corpus shows that, when engaging in conflictual interactions, Twitter users appear to make a personal decision about whether or not to allow other participants in the interaction to express different viewpoints on the topic being discussed. Users who permit the existence of divergent opinions may be open to hearing other viewpoints and may negotiate with their opponents, but those who do not may wish to force their opponents to accept an opposing opinion. The users in the latter situation in the corpus hardly employ positive politeness to satisfy the recipient’s face as they may pose a serious damage to the recipient’s negative face. For example, the speakers in the first three of the above-mentioned four sentences are more evidently concerned with changing the recipient’s opinion than the speaker in (4). (1) lists the opinions of the speaker and recipient and (2) conveys the message that the speaker has

received the recipient's point of view. This contrasts with (4), in which the speaker explicitly acknowledges the recipient's view. This speaker indicates the sincerity of their endorsement by using the adverb "compltly" (Line 239), and clarifies that they respect the recipient's right of speech, "I have no idea if any of this makes sense but yh" (Lines 240–241). Although (3) appears to play a similar role to (2), it actually also serves as a consideration of the recipient's positive face because it is adjacent to (4) and provides a basis for the clarification of innocence in (4). Therefore, (1) and (2) function to persuade the recipient to abandon their personal views or rather to adopt the (previous) speaker's point of views, while (3) and (4) serve to protect the recipient's positive face. This suggests that the speaker's claim that they share common ground or a partnership with the recipient may not be intended to preserve the recipient's positive face. It may be rather used to weaken the speaker's coercion of the recipient, thus protecting the recipient's negative face. This explains why the majority of what appears to be positive politeness in the corpus does not serve to save the recipient's positive face. This gap between the function of establishing common ground in the corpus and in Brown and Levinson's (1987/1978) study is very likely to be due to the differences between conflictual and harmonious interactions, which will be discussed in more detail in Section 7.3.

The corpus shows that although less than 0.5% (56) of the posts contain positive politeness, nearly 90% (50) of these positive politeness modifiers bring the interpersonal conflict between the speaker and the recipient down. For example, in Excerpt 109, when opposing User D34, User C34 twice tries to protect the positive face of User D34. They express successively their concerns about User D34's newly issued comments and their recognition of User D34's view, "ok honestly I get what youre saying" (Line 238) and "i[Face with tears of joy]agree compltly with what youre saying" (Lines 239–240). User D34 does not continue to argue in response. Instead, they thank User C34 for listening, "Glad that you heard my point" (Lines 242–243). They also pay attention to protecting User C34's positive face as they treat User C34 as a co-operator by showing repeatedly their understanding of User C34's viewpoint, "Sure, I get that. I get what you are saying" (Line 242). User D34's response suggests that User C34 is successful in saving User D34's face.

Excerpt 109 (Interaction 34)

Line	User Account	Text	Level	Turn
238	User C34	[1R]ok honestly I get what youre saying but I think some really did take it too literal. I[Face with tears of joy]agree compltly with what youre saying but I also agree with what I am getting at..i have no idea if any of this makes sense but yh	22nd	66
239				
240				
241				
242	User D34	[1R]Sure, I get that. I get what you are saying. Glad that you heard my point	23rd	67
243				

Another case in point is Excerpt 110. This interaction is elicited by User A35's complaint about the UK government's regulation regarding wearing masks while shopping. User A35 (Lines 153–155) acknowledges that the previous speaker (i.e., User F35) has to wear a mask due to their special job. They then oppose the mandatory mask wearing, saying that "this whole business of going mandatory for everyone is unnecessary" (Lines 155–156), and further explain this through the example of individuals with "breathing difficulties" (Lines 156–157). To respond to the objection, User F35 first highly approves User A35's view on the exemption, "I completely understand if someone is exempt from wearing a mask" (Lines 158–159), thus protecting User A35's positive face. This is followed by User F35's clarification of their own opinion. They argue that what they are really concerned with is the safety of others rather than merely mask wearing, "they take every other precaution necessary to make sure they don't have it/spread it" (Lines 159–160). Only then, in the second reply (Lines 163–167), do they begin to address their objection. User F35 does not directly criticise User A35's undifferentiated boycott of mask wearing but implies it by discussing the usefulness of wearing a mask repeatedly. They argue that for the well-being of others, people who can wear a mask should do so, "If you can wear a mask though you certainly should" (Line 163). They (Lines 165–165) then argue that they have always cared about the health of others through their experience. This point is emphasised at the end, "I want to keep the people I work for safe so I will" (Lines 166–167). By repeatedly mentioning how much they value others' health, User F35 implies an attack on User A35's selfishness. However, because their attack is indirect and ambiguous, User A35 only recognises the positive politeness used by User F35 in the first 10th level reply. They terminate the argument, thank User F35 for their respect and express appreciation for User F35's dedication.

Excerpt 110 (Interaction 35)

Line	User Account	Text	Level	Turn
153 154 155 156 157	User A35	[1R]You work in a care home, so I have no issue with you wearing one in there, you are dealing with the most vulnerable, so you can't afford to take risks. But this whole business of going mandatory for everyone is unnecessary. People with breathing difficulties are exempted.	9th	58_2
158 159 160 161 162	User F35	[1R]I completely understand if someone is exempt from wearing a mask I just hope they take every other precaution necessary to make sure they don't have it/spread it. Get tested, wash your hands, don't cough on anything and if you have any symptoms self isolate.	10th	61_1
163 164 165 166 167	User F35	[1R]If you can wear a mask though you certainly should. I haven't been wearing one in shops until now because I got tested at work and I've been negative this whole time. I don't know why they're saying it's mandatory now but I want to keep the people I work for safe so I will	10th	61_2
168 169 170 171	User A35	[1R]Respect, I'm glad we can have a conversation without insulting one another. Your work looking after the elderly often goes under appreciated, and my Grandma relies on people like yourself, so thanks for all the good work you've been doing.	11th	62

User D34's and User A35's responses also suggest that positive politeness is very useful in mitigating or even controlling personal conflict. The use of positive politeness implies an active concession on the part of speakers. Instead of insisting on assimilating the recipient's point of view and announcing the dominance of their own views, the speaker allows the recipient to have different opinions and appears to give value to this difference. Positive politeness is thus the most effective type of linguistic politeness in dealing with the conflict between the speaker and the recipient. Its influence on the overall interaction is however limited. For example, although Users C34 and D34 do cease to argue after the interaction shown in Excerpt 109, shortly afterwards, a new user joins this interaction and continues to oppose User D34 at the next level. In terms of Excerpt 110, Users A35 and F35 do stop arguing and no new user enters into their interpersonal interaction. However, the interaction between Users A35 and F35 is only part of the interaction that derives from Users A35's 1st level reply. This large-scale interaction around User A35 still continues because other users continue to criticise Users A35's boycott of mask wearing.

7.2.4 Doing conflict on-record with negative politeness

Negative politeness strategies refer to redressive actions that are suitable for reducing threats to the recipient's negative face, such as minimising the imposition on the recipient, being indirect or pessimistic, asking questions, hedging, suggesting, apologising and nominalising (Brown & Levinson, 1987/1978). The review of disagreement studies based

on face-to-face interactions found that the speaker tends to mitigate the damage to the recipient's negative face arising from disagreements by employing negative politeness such as opposing in the form of a question, adding the preface like *perhaps* and *I think/don't know*, using verbs of uncertainty and personalising arguments (e.g., Muntigl & Turnbull, 1998; Rees-Miller, 2000; Locher, 2004; Ben-Menachem & Livnat, 2018). A large amount of negative politeness, which is used to mitigate conflictual Twitter interactions, was also identified in the corpus. This study classified instances of negative politeness in the corpus into four categories based on its specific functions.

The first type involves the speaker maintaining the recipient's negative face and their own image by personalising their own point of view or weakening the force of their opposition. Typical methods include: adding hedges (e.g., *I think/believe/guess*, *in my opinion [IMO]*, *seems to me*, *probably/perhaps*), being pessimistic (e.g., *would*, *might*), asking for reasons, showing the possibility of the speaker's mistakes (e.g., *#imightbewrong*, the *if* clause) and using token agreements. It is also common for the speaker to further enhance the level of protection of the recipient's negative face by using adverbs or adverbial phrases (e.g., 'to be honest, I think that…'; 'I sincerely believe that…'). The aim is to avoid their oppositions being taken as provocative, such as User D36's personalisation in Turn 19 in Excerpt 111. Specifically, User C36 (Lines 53–56) compares the differences between men and women in three aspects: (1) length of sentence, (2) custody and (3) public concerns. They oppose the previous speaker (i.e., User D36) by arguing that men are also vulnerable in many cases, which is then challenged by User D36. User D36 (Lines 57–59) pursues the details of User C36's viewpoint, including the specific charges and the scenario and time of the occurrence. As compared to overt negativity, questioning gives User C36 an opportunity to re-evaluate the issues in their opinions, thus protecting the negative face of User C36. User D36 also personalises their objections with the words "I'm especially curious about" (Lines 57–58) and further indicates that their question is not a rhetorical question and that they expect User C36's answer through the adverb "especially" (Line 58). Although User D36's negative politeness does not immediately alleviate their personal conflict with User C36, it at least maintains the status quo of the conflict and avoids a sudden escalation. The answer given by User C36 (Lines 60–62) suggests that women are the privileged group. They also protect User D36's negative face by weakening their affirmation by using "more likely" (Line 60). At the end, they question User D36's understanding of the "Duluth Model" (Line

62), which has been used to justify the privileges of women in their last turn.[28] Rather than seeking an answer, this question aims to provide a hint for User D36 about User C36's opposition. This indirect approach also reveals User C36's greater knowledge and reflects User C36's consideration of User D36's face, as opposed to an overt objection.

Excerpt 111 (Interaction 36)

Line	User Account	Text	Level	Turn
53 54 55 56	User C36	[1R]Women get lighter sentences for the same crime. Women are more likely to get custody of the kids. When it comes down to rape, a woman victim is more likely to be taken seriously than man victim. And also the Duluth Model.	17th	18
57 58 59	User D36	[1R]What crimes do they get lighter sentences? And I'm especially curious about when/how women are more likely to be taken seriously than men?	18th	19
60 61 62	User C36	[1R]Basically for any crime. And they are more likely to be taken seriously as rape/domestic abuse victims than man victims are. Ever heard of the Duluth Model?	19th	20

The speaker also takes the recipient as the subject of the sentence in conflictual interactions, thus publicly criticising the recipient's mistakes or directly requesting or demanding the recipient. In this case, the speaker would weaken the coercion on the recipient by adding adverbs (e.g., *probably*, *possibly*), verb phrases (e.g., *seem to*, *appear to*) and modal verbs (e.g., *may*, *might*, *could*) that express uncertainty. Sometimes they may also insert hedges that are conventionally polite (e.g., *please*), add an *if* clause or indicate a negotiating tone to their sentence (e.g., using *how about*, suggesting the recipient has a try). A case in point is User C37's *if* clauses in Turn 9 in Excerpt 112. Since User D37 (Lines 18–19) objects to the conflation of men and male criminals, User C37 responds with a clarification. In the omitted Turns 6 to 8, User C37 discusses various inconveniences that men pose to women in detail and summarises feminist perspective in Turn 9 (Lines 34–38). They publicly reject User D37's statement, "No one is saying 'all men do this'" (Line 34). They further explain that in their view, the noun *man* is used as a general term, referring to the entire group of men, and expresses explicitly feminist claims. User C37 also constantly emphasises feminists' anger at men through the capitalisation of "MOST" (Line 34), the "right now" (Line 36) in a separate sentence and the verb to "deserve" (Line 37). Finally, they request that User D37 listens without provocation, "If you have ideas, awesome! If not, just listen. Hear us" (Lines 37–38). It is evident that User C37 offers User D37 the opinion of listening or

[28] Duluth model is a program developed to reduce domestic violence against women (Access, 17th January, 2022: https://www.cebc4cw.org/program/domestic-abuse-intervention-project/detailed).

leaving, rather than listening or expressing. Thus, User D37's right to express opinions is actually denied. User C37's requests actually constitute an invisible personal attack on User D37. However, in the meantime, they also attempt to protect User D37's negative face by adding *if* clauses in their request, "If you have ideas" and "If not" (Line 37). In response, User D37 (Lines 39–41) claims their common ground with User C37 on purpose, implying that their argument is justified. These are two possible reasons for User D37's use of positive politeness: (1) User D37 may interpret User C37's negative politeness in Turn 9 as face-saving or may intend solely to proactively manage their interpersonal conflict; (2) they attempt to avoid User C37 posing more blatant and serious damage to their face in the future. In the second case, User C37's turn may still be taken as a serious personal attack from User D37's perspective.

Excerpt 112 (Interaction 37)

Line	User Account	Text	Level	Turn
18 19	User D37	[1R]But that's our point. Us 99% hate that 1% too and want to stand with you against them, not be conflated with them.	4th	5
Omit Lines 20–33				
34 35 36 37 38	User C37	[1R]No one is saying "all men do this". We are saying "MOST women have experienced this at the hands of *a* man, and it needs to stop. Right now. Whatever it takes, this has to end. We deserve to not live in fear." If you have ideas, awesome! If not, just listen. Hear us.	8th	9
39 40 41	User D37	[1R]And you'll get not argument from me on that. I want to help be part of the solution, not treated as part of the problem. I don't feel like we're too far apart on this.	9th	10

The third type refers to Twitter users obscuring or removing their role in posing or receiving attacks and opposing views. A common approach is to change or remove the references to the speaker and the recipient, such as deleting the personal pronoun *I* and *you*, using indefinite pronouns or nouns (e.g., *someone*, *one*, *guys*) and using expressions like 'it appears/seems that' or 'it is acknowledged/regarded that'. As an example, in Excerpt 113, the assumption of men's curfews by the previous speaker (i.e., @Jessica) elicits User A38's objection. User A38 first describes Turn 2 as a "friendly reminder" (Line 05), providing @Jessica with the option of rejecting their rebuttal. They also further lessen the damage of @Jessica's negative face by removing all references to themself and @Jessica as the speaker and the recipient, respectively. They (Lines 05–06) then publicly oppose @Jessica, arguing that most men, including themself, behave properly. After a token agreement, "As great as all of this sounds" (Line 06), they (Lines 07–09) criticise @Jessica in an off-record manner. They compare the action of men and @Jessica and emphasise this contrast by using the

pronouns, “us” (Lines 06–07) and “you” (Line 06), thus suggesting the innocence of men and the provocation of @Jessica. User A38’s criticism completely enrages B38. Although User A38’s response is addressed to @Jessica, User B38 acts in the place of @Jessica and assumes the role of the recipient of User A38. Imitating User A38, User B38 begins the response with the sentence “Just a friendly reminder…” (Line 09), which appears to protect User A38’s negative face. User B38, however, raises a rhetorical question, “why don’t you pipe down and read the replies?” (Lines 09–10), to criticise User A38 for being provocative in an off-record manner. This suggests that the reminder is used in conjunction with the rhetorical question to denote the rudeness of User A38. User B38 then explicitly points out that User A38 “completely [miss] the point of the post” (Lines 10–11). They further reinforce the satire of User A38’s nonsense by using a patronising phrase, “well done” (Line 10), and a GIF of a woman rolling her eyes (see Figure 30).

Excerpt 113 (Interaction 38)

Line	User Account	Text	Level	Turn
05 06 07 08	User A38	[1R]Just a friendly reminder.... The majority of guys are not rapists. As great as all of this sounds, let’s not forget that you’re lumping us all into a nasty group that most of us want/have nothing to do with.	1st	2
09 10 11 12	User B38	[1R]Just a friendly reminder...why don’t you pipe down and read the replies? [2R:ER] [3R]Well done on completely missing the point of the post [4R][GIF: https://twitter.com/i/status/1047455628929773568]	2nd	3

Figure 30

The GIF in User B38’s 2nd level reply

The last type protects the recipient’s negative face and their own image by proactively explaining the reasons for the objection. This is also the most frequently used type of negative politeness in the corpus. This particular method mainly involves the opponent giving their own experience or quoting some external evidence, such as @Jessica’s Turn 5 in Excerpt 114. User A39 criticises @Jessica, the author of the original tweet, for being too negative on gender issues, “Why so negative” (Line 07). To respond to User A39’s

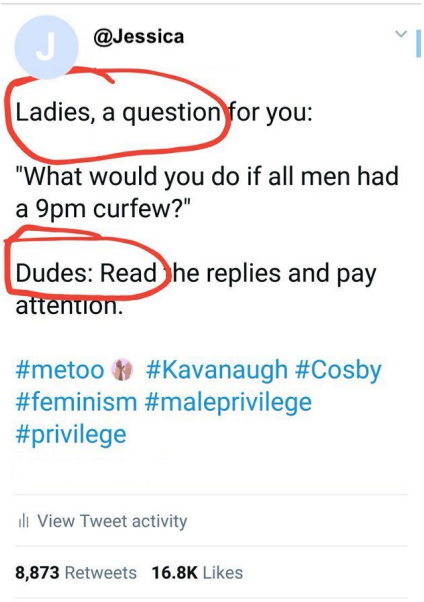
challenge, @Jessica justifies her anger at men by citing her own experiences (Lines 08–10), a screenshot of her original tweet as evidence (see Figure 31), thus protecting User A39's negative face. However, @Jessica also proactively inflames the conflict, while protecting User A39's negative face. She emphasises her annoyance and sarcasm towards men through the phrases, "fucking exhausted", "thousands of" and "entitled dude bros" (Lines 08–09). She also overtly argues that User A39 does not qualify to participate in the discussion, "This thread is not for you to speak" (Lines 10–11), and requests User A39 to listen to and respect women's opinions. Although the request is worded as an *if* clause, it actually functions to attack User A39 off-record because the options provided by the *if* clause implies a denial of the man's right to discourse. This seemingly contradictory expression suggests that @Jessica wants to deepen her critique of men while avoiding being accused of being rude. Since this phenomenon is rare in the corpus (roughly 15 replies), it is impossible to identify some common features of this phenomenon. However, it does suggest that there are some Twitter users who, even in the face of anger, are still habitually mindful of their language etiquette, just as they would be in face-to-face interactions. Since @Jessica shows a clear move to actively inflame the conflict, her Turn 5 is easily taken as a personal attack by User B39, no matter how repeatedly @Jessica uses negative politeness. User B39 (Lines 15–18) overtly questions @Jessica for being hypocritical and provocative and challenges @Jessica for blatantly inciting the conflict between genders.

Excerpt 114 (Interaction 39)

Line	User Account	Text	Level	Turn
07	User A39	[1R]Why so negative	3rd	4
08 09 10 11 12 13 14	@Jessica	[1R]Honestly because it's 4:30 in the morning and I am fucking exhausted from blocking literally thousands of entitled dude bros on this thread. [2R:ER] [3R]This thread is not for you to speak. It is for you to listen. And if you're not up for it, then we don't need you here. [4R][PIC: https://twitter.com/DanielleMuscato/status/1047780160446103552/photo/1]	4th	5
15 16 17 18	User B39	[1R]So I feminist who wants EQUAL rights yet are telling someone of the opposite sex he has NO RIGHT to comment on a PUBLIC platform?? You do realise not all men are creeps and attackers	5th	6

Figure 31

The picture in @Jessica's 4th level reply



User A38's Turn 2 in Excerpt 113 and @Jessica's Turn 5 in Excerpt 114 show that, in the corpus, the recipient rarely feels that their negative face is being protected by the speaker's negative politeness modifiers. A further example can be found in User D34's Turn 70 in Excerpt 115. User I34 joins the discussion between Users D34 and C34, which was shown in Excerpt 109, after the conflict between Users D34 and C34 has been mitigated. User I34 opposes the previous speaker (i.e., User D34), arguing that User D34 misinterprets the term "male curfew" in the original tweet. They (Lines 244–248) explain that the men's curfew is a hypothesis used to illustrate the negative impact men have on women's daily lives. User I34's explanation, however, does not convince User D34. User D34 (Lines 252–254) challenges the hypothesis, arguing that it implies a hatred of men, but their language use also highlights their preservation of User I34's negative face. The sentence "listen, that is nice" (Line 252) shows their attention towards User I34's words and their respect for User I34's response: User D34 does not just close their ears to force User I34 to accept an opposing viewpoint. In the meantime, User D34 also personalises their viewpoint by using the hedge "to me" (Lines 252–253) twice. In response, User I34 (Lines 259–267) indicates that User D34's action reflects merely man's privilege. They (Lines 259–262) attack User D34, criticising User D34 for being selfish and requesting that User D34 needs to learn compassion. However, they also protect User D34's negative face, by indicating that User D34 misinterprets their viewpoint, "if you think that's the message" (Line 259). This *if* clause, along with an adverb, "perhaps" (Line 259), further undermines their assertion. In the

second reply, User I34 escalates the attack on User D34. They (Lines 263–267) emphasise the selfishness and greed of User I34 through a comparison. This is an off-record way to do conflict and will be discussed later (see Section 7.2.5). They then blatantly attack User D34, "You're the problem" (Line 267), and also criticise User D34 for not being mature by calling User D34 "son" (Line 267). User I34's response, therefore, suggests the failure of User D34's negative politeness.

Excerpt 115 (Interaction 34)

Line	User Account	Text	Level	Turn
244 245 246 247 248	User I34	[1R]She didn't propose a curfew on all men, I hypothetical scenario- essentially she asked women what they would do differently if they knew men wouldn't be out on the streets after 9. The point being to illuminate the ways male presence harms or affects women's everyday lives.	24th	68
Omit Lines 249–251				
252 253 254	User D34	[1R]Listen, that is nice, but to me it speaks about misandry present in these kinds of discussions, the message seems to me – only if there are no men, women will finally be or feel safe.	25th	70
Omit Lines 255–258				
259 260 261 262	User I34	[1R]And btw, if you think that's the message, perhaps reflect on why that might be, and listen to what women are saying about how unsafe they feel,[1WS] instead off writing it off instantly in your head.	26th	72_1
263 264 265 266 267	User I34	[1R]So we're talking about the existential threat men pose to women every day, and the huge psychological impact that has on women, and the material effect it has on their lives, & your response is to ignore that 100% and say "who will think of the men?" You're the problem, son.	26th	72_2

The corpus also suggests that impersonalisation in conflictual Twitter interactions may also obscure the meaning of the entire sentence due to its unclear referents. In this case, impersonalisation has more in common with the practice of doing conflict in an off-record manner, as is shown in User A38's impersonalisation in Turn 2 in Excerpt 113 (also shown in Excerpt 116) and User D38's Turn 8 in Excerpt 116. The analysis of Excerpt 113 suggests that although User A38 personalises their attack, "a friendly reminder" (Line 05), User B38 (Lines 09–12) is still criticising User A38 for being provocative. This contrasts with User C38's response to User A38 in Excerpt 116. User C38 (Lines 18–19) does not interpret User A38's turn as an attack. They just clarify that women clearly know the difference between men and male criminals and also protect User A38's negative face by removing the reference to User A38 as the opponent. This suggests that User A38's impersonalisation results in an indirect and ambiguous attack. User C38's negative politeness, however, is blatantly challenged by User D38, "Your point?" (Line 20). User D38 (Lines 21–22) criticises

User C38's indiscriminate boycott of men for being a slur against themself and attacks User C38 as a man hater. Instead of referring directly to User C38 and their language, User D38 refers to "people [who] may have thought they would be raped by walking past me" (Lines 21–22) and "the language being used on this thread" (Line 20), respectively. The scope of User D38's reference is thus expanded, and the meaning of Turn 8 also becomes ambiguous like User A38's Turn 2. So the impersonalisation used by User A38 and User D38 may also serve as an off-record approach. This usage will be discussed further in the next section on doing conflict off-record.

Excerpt 116 (Interaction 38)

Line	User Account	Text	Level	Turn
05 06 07 08	User A38	[1R]Just a friendly reminder.... The majority of guys are not rapists. As great as all of this sounds, let's not forget that you're lumping us all into a nasty group that most of us want/have nothing to do with.	1st	2
Omit Lines 09–17				
18 19	User C38	[1R]Not lumping you "all" into one group. Knowing we're usually smaller & weaker than the bad guys.	2nd	7
20 21 22	User D38	[1R]Your point? The language being used on this thread is too generalised, and I am disgusted to know that people may have thought they would be raped by walking past me	3rd	8

7.2.5 Doing conflict off-record

Going off-record to do conflictual interactions means the speaker avoids overt objections/attacks and instead indicates their true intent by giving hints. According to Brown and Levinson (1987/1978), by obscuring their true intent, the speaker avoids the responsibility for confrontation, allowing the recipient to decide whether to interpret the speaker's turn as face-saving or face-threats. Although the speaker's protection of the recipient's face is not apparent, doing FTAs off-record is still considered as a politeness strategy by Brown and Levinson (1987/1978). They believed that the off-record manner suggests a relatively more evident attempt at face-saving than the on-record approach. However, in practice it is difficult to discern whether the speaker's indirectness is meant to avoid responsibility or to be polite. They added that doing FTAs off-record is considered polite only when the speaker intends to be polite and inserted brackets in "off-record (politeness)" (Brown & Levinson, 1987/1978, p. 2). Popular approaches in face-to-face interactions involve (1) being vague and ambiguous (e.g., being incomplete, using ellipses and overgeneralising) and (2) inviting conversational implications through hints (e.g., giving hints and association clues, being ironic, using metaphors and rhetorical questions) (Brown

and Levinson, 1987/1978). Both methods are also common in conflictual Twitter interactions and are described below.

The analysis shows that Twitter users typically mimic different types of laughter in conflictual interactions as a means of being ambiguous. Laughter appears in approximately 5% (555) of the posts in the corpus. It is achieved through various textual or multimodal approaches such as emoji or emoticons that mimic different types of smiling faces (e.g., 🙂, 😆, 😂), images or GIFs that depict smiling faces and body movements of humans or anime characters and texts that describe the sound of laughing or the state of smile (e.g., *hhhhhh, lol*). This approach may be due to two possible reasons. First, Twitter users are able to communicate their emotions more accurately and vividly by using laughter. Second, taking this approach, they can also conceal their real communicative intent because the emotion and/or action of laughing involves multiple types such as smiles, laughs, sneers and bitter smiles, and thus has different meanings such as friendliness, happiness, mockery and sarcasm. The recipient has to determine whether the speaker's words are meant to be friendly and polite, to reduce tensions or to be provocative and sarcastic. The use of emoji, "😂" (Line 16) and "lol" (Line 15) by User B40 in Excerpt 117 is a case in point. The previous speaker (User A40) suggests that @Jessica is supposed to substitute "a curfew for all violent offenders" for the term "male curfew". @Jessica (Lines 09–13) opposes this suggestion and argues that men hurt women not only by committing crimes but also by their daily indiscretions and obscene thoughts. In response, User B40 blatantly criticises @Jessica's idea for being "delusional" (Line 14). They (Lines 14–16) argue that the men's curfew, instead of eliminating sex crimes, would put women in a more dangerous position. User B40 also mitigates the damage to @Jessica's negative face by using an *if* clause (Line 14) and several conditional modal verbs, "would" (Lines 14–15). Meanwhile two linguistic elements that imitate belly laughs are also employed: (1) a emoji of face with tears of joy, "😂" (Line 16), and (2) an abbreviation that stands for 'laugh out loud', "lol" (Line 15). By doing so, User B40 leaves it up to @Jessica to interpret Turn 4 as a joke or irony.[29] @Jessica does not respond. User A40, who has opposed @Jessica in previous turns, however, rejects User B40's comment. There are two possibilities for the relationships among @Jessica, User A40 and

[29] Joke is a "basic positive-politeness technique, for putting (the recipient) 'at ease'" (Brown & Levinson, 1987/1978, p. 124). However, irony refers to a technique when the speaker "must cause offence, at least do so in a way which doesn't overly conflict with (being polite)" (Leech, 2014/1983, p. 233). The speaker's purpose is to allow the recipient "to arrive at the offensive point of [their] remarking indirectly, by way of implicature" (Leech, 2014/1983, p. 233).

User B40: (1) both User A40 and @Jessica are feminists and jointly oppose User B40 who considers @Jessica's view on men's curfews to be too radical; (2)although User A40 argues that User B40 is off-topic, they both oppose @Jessica who supports feminism. Both cases show that the divergence of opinions occurs equally within each of the opinion camps that endorse or disapprove of the central argument. To respond to User B40's objection, User A40 (Line 17) employs negative politeness, suggesting that User A40 does not interpret User B40's Turn 4 as a serious FTA. They personalise their view, further remove the reference to themself and leave the verb to "Think" (Line 17). They also employ the modal verb, "may" (Line 17), to provide User B40 with the option of rejecting the objection. Their use of the verb to "miss" (Line 17) is much less damaging to the opponent's face than User B40's verb to *delude*. User A40's face-saving suggests that User B40's laughter may ease the tension in a humorous manner.

Excerpt 117 (Interaction 40)

Line	User Account	Text	Level	Turn
09 10 11 12 13	@Jessica	[1R]It's not just about rape and murder. [2R:ER] [3R]It's about utter freedom from creeps, Nice Guys™, dude bros, misogynists, & mansplainers. The feeling when you & your friends are in the club bathroom & you can say what you REALLY think, you can be YOU—but anywhere, & everywhere.	2nd	3
14 15 16	User B40	[1R]Your delusional if you think [1WS]criminals would follow a curfew, all you would do is get rid of witnesses lol crim would skyrocket[Face with tears of joy]	3rd	4
17	User A40	[1R]Think you may have missed the point here Adam..	4th	5

When speakers invite implicature, they tend to suddenly utter something that appears to be irrelevant to the context and provide some hints for the implication. Comparison of two diametrically opposed matters is a common approach for cueing, such as contrasting between what the speaker supports and what the opponent approves. For example, in Excerpt 118, User I34 juxtaposes women's criticism of male privilege with men's disrespect for women, "the existential threat men pose to women every day, and the huge psychological impact that has on women, and the material effect it has on their lives" (Lines 263–265). Using this comparison, they indicate that User D34 is selfish and thus attacks User D34 off-record. User I34, subsequently announces their attack by overtly offending User D34, "You're the problem, son" (Line 267). User D34 rejects User I34's accusation, arguing that User I34 is the person who is gender-biased, "No,the problem is you" (Line 268). They (Lines 269–271) attribute the issue of User I34's viewpoint to over-generalisation and request that User I34 distinguishes between men and male criminals. They end the

conversation with a threat, “But if you keep saying all men,there wnt of a talk” (Lines 271–272). User D34’s reaction suggests that User I34’s off-record approach does not serve as face-saving. Even so, User D34 still attempts to protect User I34’s negative face because they designate User I34 and themself jointly as the addressee of the request, “We need to talk and see which men attack women and why” (Lines 270–271). This further demonstrates that users do not completely abandon conventional language etiquette in doing conflict, although uninhibited linguistic behaviour is prevalent on Twitter.

Excerpt 118 (Interaction 34)

Line	User Account	Text	Level	Turn
263 264 265 266 267	User I34	[1R]So we’re talking about the existential threat men pose to women every day, and the huge psychological impact that has on women, and the material effect it has on their lives, & your response is to ignore that 100% and say “who will think of the men?” You’re the problem, son.	26th	72_2
268 269 270 271 272	User D34	[1R]No,the problem is you. You intentionally keep saying all men pose threat to women, but that is simply not reality.Not all husbands, brothers,men in general, attack women. We need to talk and see which men attack women and why.But if you keep saying all men,there wnt of a talk.	27th	73

Hints in the corpus can be discovered occasionally through a phrase/word whose literal meaning apparently contradicts what the speaker may say rationally in the context, such as User B41’s Turn 3 in Excerpt 119. User A41 overtly opposes @Jessica by refusing to “follow said curfew” (Line 05). User B41 responds to this with “yeahhhhhhhh man, this shit isn’t about you” (Line 06), which appears to support User A41. However, by repeating the letter <h>, they simulate prolongation of the pronunciation of the word *yeah*, suggesting a strongly sarcastic and provocative tone. This tone is further stressed by the noun “shit” (Line 06), which represents the unpleasant and unfair treatment that women face. They also use the pronoun “you” (Line 06) to declare their identity as an opponent of User A41, which obviously contradicts their literal claim. User B41, thus, attacks User A41 in an off-record manner.

Excerpt 119 (Interaction 41)

Line	User Account	Text	Level	Turn
01 02 03 04	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]”What would you do if all men had a 9pm curfew?” [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
05	User A41	[1R]As a guy I wouldn’t follow said curfew.	1st	2
06	User B41	[1R]yeahhhhhhhh man, this shit isn’t about you.	2nd	3

Impersonalisation, which was discussed in Section 7.2.4, is a particular type of inviting implicature. The use of indefinite expressions obscures the speaker's reference to the involvement and viewpoints of the recipient and/or themself, whereas it is obviously cued by the explicit response relationship in interpersonal conflictual interactions. The recipient knows exactly who the speaker's intended recipient is based on whom the speaker replies to. For example, the discussion of Excerpt 116 shows that User D38 overtly challenges the previous speaker's (i.e., User C38) view and criticises User C38 for being a man hater in an off-record way. They expand the reference by using two indefinite expressions instead, namely "people [who] may have thought they would be raped by walking past me" (Lines 21–22) and "The language being used on this thread" (Line 20). User E38, who holds a similar viewpoint to User C38, takes over from User C38 and responds to User D38, as shown in Excerpt 120. They defend feminists and reject User D38's criticism of feminist opinions, "Not all men are rapists, no" (Line 23). They (Lines 23–26) further argue that men do pose great psychological pressure to women and serious threats to women's safety. User E38's response seems to announce the success of User D38's 'off-record' politeness. Their last sentence, "Small beer for you but would make a big difference to her" (Lines 25–26), however, suggests that User E38 is attempting to avoid being involved in the interpersonal conflict between Users D38 and C38 because they are well aware that User D38's turn is designed for User C38. They use the pronoun "her" (Line 26), rather than the pronoun *us*, to weaken their alignment with C38. Although this choice may be due to User E38's male identity, it is, probably more importantly, an indication that User E38 regards themself as a bystander of Users D38's interaction with User C38. Although User D38 aims to attack all feminists who share User C38's view, User D38's impersonalisation allows User E38 to refuse to acknowledge that they are the addressee of the attack. By referring to User C38, User D38 identifies themself as a bystander and distances themself from the interaction between Users C38 and D38. This suggests that, as in face-to-face interactions, 'off-record' politeness in conflictual interactions is suitable for protecting the recipient's face and preventing the speaker from taking responsibility for doing conflict. It may be also used to avoid getting involved in personal strikes, which is typical of conflictual interactions.

Excerpt 120 (Interaction 38)

Line	User Account	Text	Level	Turn
20 21 22	User D38	[1R]Your point? The language being used on this thread is too generalised, and I am disgusted to know that people may have thought they would be raped by walking past me	3rd	8
23 24 25 26	User E38	[1R]Not all men are rapists, no, but all men have the potential to be rapists in the mind of a woman walking alone in the dark. Hang back to a safe distance, or cross the road to overtake. Small beer for you but would make a big difference to her.	4th	9

'Off-record' politeness, like negative politeness, hardly makes the recipient feel that their face is being protected in the corpus, such as User I34's Turn 72 in Excerpt 118. A similar case was found in the questionnaire.[30] Participants were asked to interpret User A32's Turn 2 in Excerpt 106, which is presented here as Excerpt 121. User A32 attacks the previous speaker (i.e., @Jessica) bald-on-record, criticising @Jessica for being "condescending and shallow" (Line 07). They also protect @Jessica's face by replacing the blatant denial with a rhetorical question and wording the request as a suggestion. Based on their reports, however, around 56% (15) of the participants (27) who opposed User A32 still believed that User A32 is impolite and/or offensive due to that overt attack. This is largely because the use of negative and 'off-record' politeness does not change the nature of the speaker's utterance as an FTA. The speaker does not give up on persuading the recipient to accept an opposing viewpoint due to the use of these two types of linguistic politeness. This action is still very likely to be perceived by the recipient as an interference with their freedom of thought. To defend their arguments, the recipient accordingly may use more oppressive expressions, which explains, in part, why Twitter interactions tend to become more intense with time. However, only four participants blatantly attacked User A32 according to their responses. This further demonstrates that people are somehow influenced by conventional linguistic etiquette even when engaging in conflicting interactions on Twitter.

Excerpt 121 (Interaction 32)

Line	User Account	Text	Level	Turn
05 06 07 08 09	User A32	[1R]You do realize the vast majority of dudes don't bother or harrass women right? Your post comes off as incredibly condescending and shallow. How about instead of generalizing you start encouraging women and tell them how to protect themselves if they would like the option?	1st	2

[30] The questionnaire is used to investigate Twitter users' attitudes towards language use in conflictual interactions. See more about this survey design in Section 4.3 in Chapter 4. See the full version of this questionnaire in Appendix 2: Questionnaire 2.

7.2.6 Pseudo-politeness

The previous sections discussed how the speaker mitigates the conflict with the recipient by using positive politeness, negative politeness and 'off-record' politeness. Some posts in the corpus that appear to employ linguistic politeness are, however, used to reinforce the damage to the recipient's face. This study refers to language use that takes the form of linguistic politeness but functions to intensify conflict as *pseudo-politene*ss. For example, in Excerpt 122, User A38 argues that men are not equal to male offenders. However, User B38 believes that men, including User A38, are not qualified to have an opinion in this conversation at all. They criticise User A38 for going off-topic, "completely missing the point of the post" (Lines 10–11), but they, in turn, praise User A38 through the hedge "Well done" (Line 10). This hedge is, thus, meant to be patronising and to reinforce User B38's opposition to User A38.

Excerpt 122 (Interaction 38)

Line	User Account	Text	Level	Turn
05 06 07 08	User A38	[1R]Just a friendly reminder.... The majority of guys are not rapists. As great as all of this sounds, let's not forget that you're lumping us all into a nasty group that most of us want/have nothing to do with.	1st	2
09 10 11 12	User B38	[1R]Just a friendly reminder...why don't you pipe down and read the replies? [2R:ER] [3R]Well done on completely missing the point of the post [4R][GIF: https://twitter.com/i/status/1047455628929773568]	2nd	3

Figure 32

The GIF in User B38's 2nd level reply

Excerpt 123 is triggered by @Paul's tweet, in which he asks Twitter users what he could do to help with the Amazon fire. User A42 replies that @Paul could "film a dead body" (Line 08) in the Amazon rainforest, which constitutes a satire of @Paul, who disrespected the corpse while filming in Aokigahara, a forest in Japan. The purpose is to attack @Paul's presence as a disaster of the forest in an off-record manner. In response, User B42 mimics the laughter, "Haha" (Line 09), and seems to positively evaluate User A42's opinion, "So

funny. So original" (Line 09). The ensuing attack, however, suggests that this praise serves in fact as pseudo-politeness. User B42 overtly attacks User A42, requesting that A42 keeps silent in a quite offensive manner, "STFU" (Line 09), in order to deprive User A42's right to speak. They also camouflage their attack by adding a hedge, "How about" (Line 09). Taking this approach, they appear to offer User A42 an option. However, User B42 actually shows a strong sense of sarcasm, does not concede the right to deprive User A42's freedom of speech but further enhances their attack on User A42.

Excerpt 123 (Interaction 42)

Line	User Account	Text	Level	Turn
01 02 03 04 05 06 07	@Paul	[1R]amazon rainforest: how can I help? Need advice on how I can actually make a change here; these pictures are breaking my heart... this is one of the most important ecosystems on earth. Is there a fundraiser, a call to action, anything I can leverage my audience for on this? [2R][PIC: https://twitter.com/LoganPaul/status/1164064145198600193/photo/1]	tweet	1
08	User A42	[1R]film a dead body in it	1st	2
09	User B42	[1R]Haha. So funny. So original. How about you STFU	2nd	3

Figure 33

The picture in @Paul's tweet

Another example is User C34's Turn 10 in Excerpt 124. After User C43 and D43 debate the relationship between women's safety and men in several turns, User D43 questions User C43's purpose. They (Lines 35–36) criticise User C43 for being provocative, thus attacking User C43 off-record. User C43 interprets User D43's turn as an attack and blatantly refuses to cooperate. They provoke User D43, arguing that they would invite more Twitter users to join the opposition, "I'll follow your lead and try to tell other people to talk somewhere else" (Lines 37–38). However, they end their turn with a phrase that usually serves as a polite expression to end an email, "Thanks and good day" (Line 38). It thanks the addressee for reading the email and also offers good wishes to the addressee. However, since this phrase is located after User C43's blatant attack on User D43, it functions to

coercively terminate the interaction with User D43, which shows no sign of a closing sequence. This results in User C43 deliberately interfering with User D43's freedom of speech.

Excerpt 124 (Interaction 43)

Line	User Account	Text	Level	Turn
35 36	User D43	[1R]Do you want to be part of the solution or what? Or are you going to just come on threads like this and complain.	9th	9_2
37 38	User C43	[1R]No I'll follow your lead and try to tell other people to talk somewhere else.[1WS] Thanks and good day	10th	10

As pseudo-politeness is the only linguistic politeness in the corpus used to reinforce face-damage, it is reasonable to assume that such politeness is very common on Twitter. The analysis, however, shows that there are less than 1% (70) of the posts containing pseudo-politeness in the corpus. This study then designed a questionnaire survey for Twitter users to gain a more comprehensive understanding of this issue (see survey design in Section 4.3 in Chapter 4 and see the full version of the questionnaire in Appendix 2: Questionnaire 2). It surveyed the differences between the use of pseudo-politeness in face-to-face and Twitter contexts, mainly based on a typical form of 'Can you please fuck yourself?'. The result shows that less than 17% (5) of the participants declared using such structures on Twitter and 30% (9) in face-to-face settings, which seems to suggest that people are more likely to use pseudo-politeness in face-to-face interactions than in Twitter interactions. However, half of the participants reported that they had encountered other forms of pseudo-politeness on Twitter, such as "Please, it is none of your business, OK?" and "Come on, don't be this aggressive, be a grown-up". The data used in the present study, therefore, can neither verify the assumption nor disprove it. In terms of perception, 90% (27) of the participants consider pseudo-politeness as a type of language use employed for sarcasm and ridicule with serious consequences. This suggests that pseudo-politeness is a well-known technique for provoking the opponent in conflictual interactions. This further reflects that although pseudo-politeness may not be widely used, it is indeed typical of Twitter, which is known for the prevalence of conflictual interactions.

The review found that the linguistic research on pseudo-politeness has been undertaken especially in recent years (e.g., Culpeper, 2001; Leech, 2014/1983; Aijmer, 2015; Simonin, 2018; Graham, 2019; Murphy, 2019; see Section 3.4.3 in Chapter 3). Linguists have observed that pseudo-politeness emerges in the form of politeness and serves to damage

the recipient's face. However, to my knowledge, the interactional functions of such politeness have not been studied in much depth yet. The corpus shows that the use and interpretation of pseudo-politeness are highly affected by the concept of rationality in Brown and Levinson's work (1987/1978). As one basic assumption, rationality refers to "the capacity and interest that the attitudes, decisions, beliefs, etc. of one who has such property are well grounded or validated and that consequences can be drawn from the reasoning involved in the rational process" (Giarolo, 2018, p. 224). Brown and Levinson (1987/1978, p. 59) postulated that "all MPs are rational agents". That is, participants in an interaction possess "certain rational capacities" (Brown & Levinson, 1987/1978, p. 59), in particular, the ability to "choose means that will satisfy their ends" (Brown & Levinson, 1987/1978, p. 61). Thus, to maintain the recipient's face, the speaker can employ suitable linguistic politeness to redress their FTAs. There is, however, apparent irrationality in the three turns of Users B38, B42 and C43. The hedge "well done" (Line 10 in Excerpt 122) suggests User B38 is praising and applauding User A38, which conflicts with their blatant negation, "completely missing the point of the post" (Lines 10–11). Although the hedge "How about" (Line 09 in Excerpt 123) is used, User B42 still denies User A42's right to speak by arguing "STFU" (Line 09). The phrase "Thanks and good day" (Line 38 in Excerpt 124) normally serves to end an E-mail politely, but User C43 employs it to coercively terminate a conflictual interaction, thus interfering with User D43's freedom of speech. Users B38, B42 and C43 create logical conflict through apparently polite statements and starkly opposing positions expressed in the discourse. The linguistic politeness that they employ is not designed for minimising FTAs but rather serves as pseudo-politeness. The purpose is to increase the offensive implicature of features such as irony, sarcasm, provocation and patronisation and thus to aggravate the recipient's face injury and intensify the conflict.

In order to create this logical conflict, Twitter users are well aware that they need to maximise the contrast between polite and offensive parts of their utterance, thus avoiding linguistic politeness being perceived as face-saving. The use of conventional politeness in opposition (including both disagreements and personal attacks) with grossly offensive implicature is the most common method for creating this logical conflict in the corpus. To be "conventionally polite", people are supposed to adhere to the politeness that is governed by the "norms of politeness" (Pfister, 2010, p. 1276), whose primary rule is that a "politeness evaluation arises when an action is incongruence with a norm" (Fraser, 1990, p. 219). A few

examples include: expressing gratitude, saying hello when meeting someone, using the hedge *please* when bothering someone to do something and ending the conversation with a closing remark. Brown and Levinson (1987/1978), however, did not give special attention to conventional politeness because they argued that politeness comes from communication, rather than being intrinsic to language. This contrasts with a common belief that conventional politeness in most cases serves as a panacea for face-to-face interactions, especially harmonious ones. For example, in English-speaking cultures, using the hedge *please* when asking someone to do something is generally considered polite. This leads to the use of the hedge *please* and other forms of conventional politeness in conflictual interactions, colliding with people's default of what type of language should be used in such situations. In other words, the use of conventional politeness violates the Gricean Quality Maxim (i.e., do not say more than is needed). Due to this collision, the recipient has to seek to infer the speaker's strategic goal and not to mistake the speaker's conventional politeness as kindness and friendliness. As a result, conventional politeness becomes the most common and primary form of pseudo-politeness.

The corpus also suggests that, occasionally, linguistic politeness that does not take the form of conventional politeness (hereinafter called *unconventional politeness*), may also function as pseudo-politeness. A typical case in point is the hedge *I think*. It is normally used as negative politeness because it is suitable for protecting the recipient's negative face by personalising the opposing opinion. However, when it is used with blatant personal attacks, it may serve as pseudo-politeness. For example, to respond to the challenge from the previous speaker, User A18 (Line 33) in Excerpt 125 explains their reasons for refusing to wear a mask. User C18 (Lines 34–38) attacks User A18, criticising User A18 for having ulterior motives and being selfish (see more discussions in Excerpt 87). They also add the hedge "I think" (Line 37) to emphasise their role as the authority of the attack, thus provoking User A18. The hedge *I think* in the phrase "I think the word for someone who won't wear them in public is selfishtwat" (Lines 37–38) thus serves as pseudo-politeness.

Excerpt 125 (Interaction 18)

Line	User Account	Text	Level	Turn
33	User A18	[1R]Cos it wasn't relevant 4 months ago and isn't now	9th	10_1
34	User C18	[1R]I bet you clapped on Thursdays for the NHS doing a great job, just to disrespect them online and potentially cause a second lockdown. You can call ppl who wear masks sheep's all you want, but I think the word for someone who won't wear them in public is selfishtwat.	10th	11
35				
36				
37				
38				

However, the hedge *I think* used in personal attacks does not always function to intensify FTAs. For instance, in Excerpt 126, to answer the previous speaker's (i.e., User B44) query about mask wearing in Covid-19, User A44 (Lines 09–13) states in detail the functions of masks in different situations. User B44 rejects User A44's point by displaying the data from NHS England and blatantly attacking User A44 for having IQ problems, "your mask has suffocated your brain" (Line 26). They also add the hedge "I think" (Line 26) to the overt attack, which seems to reinforce their FTA; but, actually, User B44 intends to use it to camouflage their attack on User A44. User B44 leaves the judgement to User A44 and other recipients to decide the degree of the FTA, thus going off-record to attack User A44. The hedge *I think* in the phrase "I think your mask has suffocated your brain" (Line 26) thus serves as 'off-record' politeness. The comparison of User C18's Turn 11 in Excerpt 125 and User B44's Turn 9 in Excerpt 126 shows that unconventional politeness may play a different role when used with personal attacks. In Excerpt 126, User B44 intends to injure User A44's face without accepting the responsibility for the injury. By avoiding the use of abusive and/or discriminatory terms such as swear words, they disguise their intent to do an FTA and further conceal this intent by adding the hedge *I think*. This contrasts with Excerpt 125. In order to emphasise their deliberate provocation, User C18 employs the word "selfishtwat" (Line 38), which blatantly indicates that they are insulting User A18. They then further maximise the contrast between the politeness and conflict exchanges in their utterance by using the hedge *I think*. As such, unconventional politeness serves as pseudo-politeness only when it is applied to personal attacks containing linguistic expressions that convey an abusive and/or discriminatory implication; this implication is expressed explicitly and is hardly influenced by the context.

Excerpt 126 (Interaction 44)

Line	User Account	Text	Level	Turn
09	User A44	[1R]Hello – you're really not understanding how masks work. If someone wore a mask, but comes in contact with someone that has Covid-19, they're likely to get infected. If someone that's wearing a mask has Covid-19 comes in contact with someone that hasn't got the virus, they're...	4th	4_2
10				
11				
12				
13				
Omit Lines 14–22				
23	User B44	[1R]NHS England figures from last week indicate that 13 people in the entire country died from the virus. Overall, the IFR is 0.04% and falling, 99.96% survive and 80% have no symptoms. I think your mask has suffocated your brain.	5th	9
24				
25				
26				

7.3 Comparing the use of linguistic politeness in Twitter and face-to-face settings

The previous section primarily discussed the distribution of four types of linguistic politeness in the corpus (i.e., positive politeness, negative politeness, 'off-record' politeness and pseudo-politeness) and their strategic goals and discursive functions. It shows that only less than 0.1% (34) of the posts in the corpus that are perceived as a personal attack by the recipient (4170) consist of pseudo-politeness or are phrased bald-on-record; all of the remaining posts contain linguistic politeness modifiers that function to mitigate FTAs. However, Twitter is still stereotypically perceived to be highly aggressive. In order to provide insights into this stereotype, this section further compares the use of linguistic politeness in conflictual Twitter interactions with those described by Brown and Levinson (1987/1978), from a more holistic perspective. It argues that the formation of this stereotype is due to the following three major factors.

First, people tend to incorrectly assert that maintaining social harmony is the only goal and outcome of using linguistic politeness based on their experience in face-to-face communication, including Brown and Levinson (1987/1978). They assumed that there is a one-on-one relationship between linguistic forms/strategies and their communicative impact (see Section 3.4.3 in Chapter 3), regarding linguistic politeness as a method to secure social harmony in human communication. They then verified this assumption by examining the use of linguistic politeness in harmonious face-to-face interactions and/or face-to-face interactions designed to maintain a harmonious relationship between interlocutors. However, the present study challenges this assumption by examining linguistic politeness in conflictual Twitter interactions. It finds that the default of conflictual Twitter interactions differs from that of face-to-face interactions identified by Brown and Levinson (1987/1978). Participants tacitly regard refuting their opponents' opinions and/or attacking the

opponents, rather than seeking social harmony, as their primary task and believe that their opponents share the same perception about the task as they do. Therefore, when the speaker puts forward linguistic politeness in conflictual interactions, the recipient needs to draw an inference about whether the (previous) speaker is concealing any ulterior motives. Although in face-to-face arguments, people are able to address this issue by referring to paralanguage features such as intonation, pitch, speed of speaking, eye contact, facial expressions, gestures, postures and hesitation noises, this is not the case on Twitter. Twitter algorithms erase most of these paralanguage features but create multimodal features, including emoji, GIFs, pictures and videos. People thus have to adapt their ways of obtaining information in face-to-face contexts to the case of Twitter. They identify the speaker's literal meaning and reasonably infer the speaker's implications and communicative intents based on the discursive context.

The analysis of conflictual Twitter interactions subsequently shows that in such interactions, maintaining a harmonious relationship is not the only function of linguistic politeness nor the only reason why Twitter users employ linguistic politeness. As shown in Table 10, pseudo-politeness is used to maximise the damage to the opponent's face by blatantly announcing Twitter users' provocative or sarcastic intent, thus reinforcing the FTAs. By adding another possibility of polite intent to Twitter users' offensive intent, 'off-record' politeness is used to do FTAs in an indirect manner to disguise the FTAs. Without asking the speaker to give up their strategic goals (i.e., guaranteeing the dominance of their own viewpoint), negative politeness is employed to protect their own image and to minimise the damage caused by the speaker's FTAs to the opponent's negative face. Although both 'off-record' politeness and negative politeness project the speaker's face-protection in most cases, they still fail to convince the recipient to consider that their face has been adequately preserved and thus tend to escalate conflict to some extent. In terms of positive politeness, although it is the only type of linguistic politeness that actually serves to mitigate or control the personal conflict between interlocutors, its manifestations are also quite different from those identified by Brown and Levinson (1987/1978). For instance, they regard the approach of exaggerating interest, approval and sympathy as positive politeness in face-to-face harmonious interactions because the underlying assumption is that the speaker attempts to establish common ground with the recipient. The sentence "What a fantastic garden you have" (Brown & Levinson, 1987/1978, p. 104) is thus much more polite than the sentence

'your garden is fine' because the use of exclamation undermines the damage to the recipient's positive face by exaggerating the speaker's approval to the recipient's belongings. In conflictual Twitter interactions, exaggerated expressions are likewise used to protect the recipient's positive face, although less commonly. For example, in the sentence "I completely understand if someone is exempt from wearing a mask" (Line 158; see Excerpt 110), the adverb *completely* is used by User F35 to emphasise their support of the previous speaker (i.e., User A35). However, when the exaggerated expression is used in a blatant attack or disagreement, it rather functions to intensify the conflict. For instance, in Excerpt 114, @Jessica emphasises her annoyance and irritation of the previous speaker (i.e., User A39), by using "fucking exhausted" (Lines 08–09) and "thousands of" (Line 09). These exaggerations are not intended to protect User A39's positive face nor are they used as pseudo-politeness because they reinforce the FTAs by going on-record, rather than by maximising the contrast between @Jessica's politeness and opposition.

Table 10

The strategic goals and functions of different types of linguistic politeness modifiers in conflictual Twitter interactions

Linguistic politeness	Purpose	Impact
Pseudo-Politeness	To maximise the face-damage to the recipient;	Definitely angering the recipient and greatly intensifying the conflict;
Off-Record Politeness	To disguise the face-damage to the recipient and to protect their personal image;	Occasionally making the recipient feel face-saving and slightly intensifying the conflict;
Negative Politeness	To minimise the face-damage to the recipient and to protect their personal image;	
Positive Politeness	To give up persuading the recipient to accept an opposing viewpoint, thus greatly minimising the damage to the recipient's face;	Generally preserving the recipient's face and successfully resolving the interpersonal conflict with the recipient.

Second, even if speakers' linguistic politeness is designed to redress FTAs, it appears to be ineffective in mitigating conflict on Twitter. The 'ineffective' here does not mean that linguistic politeness objectively fails to save the recipient's face completely because the analysis shows that positive politeness is still a powerful method for mitigating the speaker's personal conflict with the recipient. It rather means that linguistic politeness in conflictual Twitter interactions is less effective in protecting the recipient's face than in face-to-face interactions. As Sections 7.2.3 and 7.2.4 suggest, even if the speaker employs negative politeness and 'off-record' politeness to preserve the recipient's face, in most cases, the recipient does not believe that the speaker's protection of their face is sufficient to offset

the speaker's damage to their face. There is no doubt that, when being attacked, the recipient is able to feel the speaker's face-protection. For example, as the discussion in Chapter 5 (see Section 5.3.1) suggests, User C2 (Lines 12–15) implies the non-existence of God by arguing that God lacks compassion, which triggers a conflictual interaction with User B2. In Excerpt 127, a new user, User E2 also challenges User C2's atheistic view and triggers another interaction between Users B2 and C2. User E2 (Lines 100–101) accuses User C2 of being cold-blooded and lacking in compassion, which is then supported by User B2 (Lines 102–104). In response to User B2, User C2 (Lines 105–106) first defends themself by refusing to accept the logical connection between being atheists and the absence of compassion. They then argue that "I firmly believe that you can be a good person by doing good deeds" (Lines 106–107), which seems to protect the positive face of User B2. This statement, however, functions as an off-record attack as User C2 intends to imply that User B2 is a person without compassion. They terminate the turn by blatantly attacking User B2 for being "arrogant & insulting" (Line 108) and preserve User B2's negative face by explaining the reason for the attack, "I am hard hearted just because I don't have a god" (Lines 108–109). Taking a similar approach, after overtly rejecting User C2's accusation, "That's NOT what anyone was saying" (Line 110), User B2 overtly attacks User C2. They challenge User C2's comprehension ability, "work on your reading comprehension" (Lines 110–111), and question User C2 for having ulterior motives, "not as dumb as you make out?" (Line 112). They then explicitly characterise User C2's behaviour as an attack, urging User C2 to "stop attacking" (Line 112). User B2 also accuses User C2 of being "weak and pathetic" (Line 113) and publicly interferes with User C2's freedom of speech through a farewell, "Ciao" (Line 113). Meanwhile, User B2 also saves User C2's face by phrasing the request as a suggestion, "I suggest you" (Line 110), and by inserting a smiley emoji "  " (Line 114) to obscure their aggressive intent. User B2's response suggests that User B2 is able to perceive User C2's face-preservation, but the mismatch still occurs between User B2's action and User C2's use of linguistic politesse. This suggests that the standard by which Twitter users judge politeness may have changed, which constitutes the third possible reason for the stereotype that Twitter interactions are rude.

Excerpt 127 (Interaction 2)

Line	User Account	Text	Level	Turn
12	User C2	[1R]I'm assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn't sound like a very compassionate deity to me.	3rd	4
13				
14				
15				
Omit Lines 15–99				
100	User E2	[1R]May Allah strengthen your mind and soften your heart. Ameen	4th	22
101				
102	User B2	[1R]That's all I was trying to tell him![1WS] "Soften your heart", was well put, brother[Smiling face with smiling eyes][Folded hands][Red heart]	5th	23
103				
104				
105	User C2	[1R]Now look here folks – just because I don't believe in a religion doesn't mean that I have no compassion. I firmly believe that you can be a good person by doing good deeds – and I do. It is arrogant & insulting to infer that I am hard hearted just because I don't have a god!	6th	24
106				
107				
108				
109				
110	User B2	[1R]That's NOT what anyone was saying.[1WS] I suggest you work on your reading comprehension, or perhaps you're just crying wolf, and not as dumb as you make out?[1WS] Either way, stop attacking people for their beliefs.... it's weak and pathetic. Ciao[Smiling face with open mouth and smiling eyes]	7th	25
111				
112				
113				
114				

Third, Twitter algorithms shape this stereotype and further alter the underlying interactional norms. Chapter 1 shows that Twitter algorithms create a virtual space where users tend to experience a sense of unreality and a sense of being separated from their daily lives. They are allowed to partially or fully conceal their real-life identities and to create new ones (Ilbury, 2020) and do not have to receive any substantial penalties that threaten their reputations, health or wealth in real life for not using language in a courteous manner (e.g., "Twitter—Hateful conduct policy", 2022; "Twitter—Sensitive media policy", 2022; "Twitter—Violent threats policy", 2022; "Twitter—How to report content and assign moderators", 2022). This results in the prevalence of uninhibited linguistic behaviour on Twitter and also an increased level of rudeness and/or aggressiveness in arguments. There is evidence in this chapter that people do use bald-on-record oppositions and pseudo-politeness modifiers when engaging in a conflictual Twitter interaction, although these methods only appear in less than 4% (412) of the posts in the corpus. However, it is very likely that this small number of bald-on-record oppositions and pseudo-politeness approaches collide greatly with people's default of politeness in face-to-face interactions, due to two reasons. First, Twitter algorithms enable these inharmonious language uses to appear in the form of a written language on every user's screen, where users can clearly and precisely detect conflict. Second, according to Chapter 6, offensive and aggressive posts are likely to receive a large number of comments and to continue to attract new users to the discussion because of the

storage capacity of Twitter. As the amount of comments on these posts increases, these posts tend to be taken as popular by Twitter algorithms and to be given priority in the system's default post ranking. By doing so, these posts that are full of offensive and aggressive language uses may be captured by more Twitter users and may function to characterise these users' perception of Twitter. In other words, the perception of rudeness left by bald-on-record oppositions and pseudo-politeness approaches may produce a huge impact on people's perceptions, disturbing people's overall perception of Twitter. This eventually becomes a stereotype. Therefore, the threshold for the recipient to judge that the speaker's politeness in face-to-face contexts serves as face-saving is likely to be increased on Twitter. The speaker has to make a greater effort than in face-to-face environments to convince the recipient of their innocence and courtesy, such as allowing the equal presence of different opinions in the interaction. A typical case in point is the absence of the closing sequence in conflictual Twitter interactions that was identified in Chapter 5 (see Section 5.1.1).

In harmonious interactions, the closing sequence is generally considered as a linguistic courtesy to terminate the conversation overtly because it protects the interlocutor's positive face by stating or implying that the conversation is enjoyable and that the relationship can continue (Coppock, 2005). It also provides an alternative option for the interlocutor to end the conversation by avoiding making the interlocutor feel disenfranchised due to the termination. The closing sequence is not merely a matter of indicating the speaker's willingness to end an interaction. More importantly, the speaker needs to show that the termination of the conversation is negotiable by waiting for the recipient's response, which suggests the speaker's respect for the recipient. In conflictual Twitter interactions, however, the closing sequence is often used as pseudo-politeness, such as User C43's "Thanks and good day" (Line 38) in Excerpt 124. The speaker tends to express their desire to terminate as a notification or command without providing the recipient with any room for negotiation because, in their viewpoint, maintaining relationship in Twitter conflict is not necessary. They are concerned with how to broadcast their own viewpoints and to win arguments through this output. By using the closing sequence in a conflictual interaction, the speaker announces their decision to end the conversation and deprives the recipient of the right to speak. They also construct a situation that challenges the recipient's default to ensure that the recipient is able to explore the speaker's implication of anger. In

order to protect their right to speak, the recipient may ignore the speaker's desire and force the conversation to continue and is also likely to pose a severe face-threat to the (previous) speaker in turn, such as a blatant attack, to ensure that their anger is registered. In general, the closing sequence on Twitter thus does not really terminate conflictual interactions but rather is highly likely to trigger the intensification of the interaction. As Chapter 5 suggests, stopping to reply, however, becomes the most common way. An abrupt withdrawal is an efficient and convenient way for Twitter users to leave an interaction (either harmonious or conflictual) because users are not monitored to provide a response on Twitter as in face-to-face settings. Withdrawals also serve as a method of self-protection for Twitter users by offending their opponents in an unmarked manner. Unlike WhatsApp and Messenger, where messages that have been checked are marked, the only way for Twitter users to know whether the previous speaker has checked their comments is when they see the previous speaker's response. In other words, without their opponents' replies, Twitter users cannot tell whether they have successfully rejected their opponents' opinions or attacked the opponent. This may result in Twitter users feeling that they are ignored and disrespected.

7.4 Summary and conclusion

This chapter examined how the speaker uses linguistic politeness to regulate the intensity of their conflict with the recipient, based on Brown and Levinson's face theory (1987/1978). It treats linguistic politeness differently from Brown and Levinson (1987/1978), who considered it as a linguistic strategy to redress FTAs. This study, however, treats it as a modifier to adjust the intensity of conflictual Twitter interactions, including mitigating or reinforcing the FTAs.

It argues that although Twitter is generally perceived as a medium where face concerns have a low priority, users still employ linguistic politeness modifiers regularly even when engaging in conflictual interactions. As Table 11 suggests, in order to protect their personal image from taking responsibility for blatantly exacerbating conflict, the speaker overwhelmingly uses negative politeness and/or 'off-record' politeness and seldomly employs positive politeness to repair the damage caused by their disagreements and personal attacks to the recipient's face. The specific methods include: (1) personalising or weakening the strength of the objections and attacks, (2) mitigating public criticism or blatant requests by hedging or switching expression patterns, (3) concealing or removing the

authority and/or the receiver of the objections and attacks, (4) explaining the reasons for the objections and/or attacks, (5) expressing ambiguously, (6) inviting conversational implications through hints and (7) giving up the dominance of their own opinion.

Table 11

A summary of findings in Chapter 7

<table>
<tr><th colspan="3">Linguistic politeness</th><th>In face-to-face interactions[31]</th><th>In conflictual Twitter interactions</th></tr>
<tr><td rowspan="5">Mitigating FTAs</td><td rowspan="3">On-record</td><td rowspan="2">Positive politeness</td><td colspan="2"><u>Similarities:</u>
1. Manifestations;</td></tr>
<tr><td><u>Differences:</u>
1. Frequency of use: high;
2. Purpose: to save the recipient's positive face;
3. Impact: saving the recipient's positive face;</td><td><u>Differences:</u>
1. Frequency of use: low;
2. Purpose:
(1) To save the recipient's positive face;
(2) To manage the interpersonal conflict;
3. Impact: generally functioning to resolve the interpersonal conflict;</td></tr>
<tr><td rowspan="2">Negative Politeness</td><td colspan="2"><u>Similarities:</u>
1. Frequency: high;</td></tr>
<tr><td><u>Differences:</u>
1. Purpose: to save the recipient's negative face;
2. Manifestations;
3. Impact: saving the recipient's negative face;</td><td><u>Differences:</u>
1. Purpose:
(1) To save the speaker's personal image;
(2) To save the recipient's negative face;
2. Manifestations: adding what appears to be positive politeness (e.g., establishing common ground, claiming the relationship as co-operators);
3. Impact:
(1) Seldom saving the recipient's negative face;
(2) Generally slightly intensifying the conflict;</td></tr>
<tr><td rowspan="2">Off-record</td><td rowspan="2">'Off-record' politeness</td><td colspan="2"><u>Similarities:</u>
1. Frequency: high;
2. Purpose: to do FTAs without taking any responsibilities;
3. Manifestations;</td></tr>
<tr><td><u>Differences:</u>
1. Impact: occasionally saving the recipient's face;</td><td><u>Differences:</u>
1. Impact:
(1) Seldom save the recipient's face;
(2) Generally slightly intensifying the conflict;</td></tr>
</table>

(To be continued)

[31] Harmonious ones and ones that aim to establish a harmonious relationship.

<table>
<tr><td colspan="2">Linguistic politeness</td><td>In face-to-face interactions[31]</td><td>In conflictual Twitter interactions</td></tr>
<tr><td rowspan="2">Intensifying FTAs</td><td rowspan="2">Pseudo-politeness</td><td colspan="2"><u>Similarities:</u>
1. Purpose: to increase the offensive implicature (e.g., irony, sarcasm, provocation and patronisation), thus damaging the recipient's face;
2. Impact:
(1) Seriously damaging the recipient's face;
(2) Greatly intensifying the conflict;
3. Manifestations: taking the form of politeness;</td></tr>
<tr><td><u>Differences:</u>
1. Frequency: (Lacking references)</td><td><u>Differences:</u>
1. Frequency: low;</td></tr>
</table>

The formation of the stereotype that Twitter lacks courtesy was also discussed in this chapter. First, people's experiences in face-to-face interactions tend to lead them to wrongly perceive the creation and maintenance of social harmony as the only goal and outcome of linguistic politeness. Second, although positive politeness is quite effective in terminating or even resolving interpersonal conflict between interlocutors as it requires the speaker to cease to fight for the dominance of opinion, this approach is very rare in conflictual Twitter interactions. In terms of negative politeness and 'off-record' politeness that are most commonly used on Twitter, neither approach is very effective in protecting the speaker's personal image or in redressing the recipient's anger. Third, Twitter users do sometimes conduct conflict in a bald-on-record manner and/or employ pseudo-politeness to exacerbate the damage to their opponent's face. Using pseudo-politeness means that the speaker creates a logical conflict to announce their provocative and/or sarcastic intent blatantly. The speaker tends to compare what appears to be linguistic politeness with the default that maintaining relationships is much less important than broadcasting opinions when engaging in conflict. The common approaches include: using conventional politeness in highly offensive contexts and using personal attacks containing abusive or discriminatory words with unconventional politeness. Fourth, Twitter algorithms facilitate the posts that involve the bald-on-record manner and/or pseudo-politeness methods to be noticed and remembered by more users. This results in this very small number of bald-on-record oppositions and pseudo-politeness approaches producing a huge impact on people's perceptions. It disturbs people's overall perception, creates the stereotype that Twitter is rude and significantly enhances the threshold of what can be interpreted as polite on Twitter. As a result, linguistic politeness becomes extraordinarily difficult to achieve on Twitter and in turn further reinforces this stereotype of Twitter.

Chapter 8 Conclusion

8.0 Introduction

This chapter first summarises the findings of the present study (see Section 8.1). It then discusses the conclusion and outlines the implication of this study from four perspectives (see Section 8.2): (1) how can the analysis of conflictual interactions contribute to the development of pragmatic research? (2) By researching Twitter users' language practices, what relationships can be revealed between technology and natural language/human communication? (3) What methodological guidance can be drawn from the analysis of Twitter interactions for future CMC studies? (4) How can this study provide insights into the regulation of linguistic behaviour and the improvement of verbal hygiene in online environments? This chapter finally assesses the constraints of this study and provides an outlook (see Section 8.3).

8.1 Summary of findings

The widespread presence of uninhibited linguistic behaviour on the Internet threatens cyber security and social stability and has promoted the necessity for verbal hygiene in virtual environments. To address this issue, computer scientists have developed a variety of text-based computational classifiers, but these classifiers appear to have reduced effectiveness in detecting hateful and harmful language use on the Internet. Linguists have since then been concerned with the discursive nature of conflict because of this failure of linguistic features in their application. Computational technologies also make it possible for people to manifest uninhibited linguistic behaviour in written forms, thus enabling linguists to recognise and document inharmonious language phenomena that occur in virtual settings. Therefore, the present study focused on the discursive nature of online conflict, aiming to explore the developmental mechanisms of conflictual interactions in CMC (see Chapter 1).

By reviewing linguistic research on conflictual interactions, this study found that researchers are primarily concerned with two issues: (1) what linguistic strategies are used by people from a variety of social groups across diverse social contexts to oppose their opponents and (2) how conflictual linguistic strategies initiate and terminate interactions between interlocutors. Researchers mostly adopt the classic methodological approaches

that are commonly employed in pragmatic research. That is, they collect conflictual interactions which provide authentic interactional content and contextual information and analyse these interactions by using DA and (im)politeness theories. In some cases, sociolinguistic surveys are also conducted to verify the results of textual analysis and to provide additional information, particularly participants' perceptions and attitudes to the conflictual interaction.

These studies have provided a number of useful insights into different types of disagreement and personal attack strategies, the reasons and impacts for using a particular type and the distribution and characteristics of different stages in conflictual interactions. However, they have not delved deeply into people's use of pragmatic strategies based on online datasets or paid much attention to extensive forms of personal attacks other than hate speech. They have also not dealt with the role of linguistic strategies in driving the development of interactions from beginning to end without interruption and have not examined how the use of these strategies changes the intensity of conflict as the interaction evolves. A number of unique features of conflictual interactions in CMC such as their nature, length, duration, scope, structures and participants also significantly challenge the classic data collection and analysis methods employed in these studies. This is because these methods were designed for researching harmonious interactions or interactions that aim to maintain or establish harmonious relationships in face-to-face contexts.

In order to address these issues, this study decided to pay close attention to how people use linguistic strategies to initiate and progress conflictual interactions on Twitter and how people engage in conflictual Twitter interactions under the constraints and reshaping of Twitter algorithms. Three specific tasks were then identified in Chapter 2 (see Section 2.3): (1) to identify linguistic strategies for (de-)escalating conflictual Twitter interactions, (2) to explore the role of these strategies in the progress of conflictual Twitter interactions and (3) to examine the primary makeup and discursive functions of linguistic politeness in conflictual Twitter interactions. This study also integrated data collection and analysis approaches from sociolinguistics and corpus linguistics into the pragmatic analysis, establishing a pragmatics-based multidisciplinary methodological framework and conducting a sociolinguistics-oriented pragmatic analysis. Specifically, this study created a corpus of conflictual Twitter interactions containing 38 tweets and 11,188 replies with approximately 32,000 words and analysed it by using DA, Brown and Levinson's (1987/1978) facework and

some basic quantitative approaches, including frequency proportion and Pearson correlation coefficient. This study also conducted two short questionnaire surveys to collect regular Twitter users' understandings of the use of linguistic politeness and their attitudes towards Twitter conflict.

The analysis found that while the de-escalation stage has been extensively examined in linguistic research of face-to-face arguments (e.g., Vuchinich, 1990; Shum & Lee, 2013; Dobs & Blitvich, 2014) due to its importance in face-to-face contexts, this stage is particularly rare in conflictual Twitter interactions. On Twitter, users' arguments tend to terminate abruptly when one party withdraws from the conflict or to seem to terminate in certain streams but then reappear in other streams. This Twitter-specific behaviour which flouts interactional norms in face-to-face interactions may be attributed to five possible reasons that appear to be unique to CMC, or rather Twitter interactions. First, although Twitter assumes the role of the third party to monitor and manage its users' linguistic behaviour, its control over conflict such as suspending accounts and deleting posts, tends to occur after conflictual interactions have ended. Second, since the management of conflict by the Twitter company and government institutions does not substantially damage the reputations, relationships, social status or physical health of Twitter users in the real world, users' language uses are not constrained by the pressure of being penalised for observing or participating in conflict. Third, Twitter does not notify the speaker whether or not their reply has been read by the recipient, as WhatsApp and Messenger do. This function enables Twitter users to pretend that they have not received the message and they do not worry about being accused of lacking courtesy for withdrawing from the interaction. Fourth, since the Twitter company allows users to communicate anonymously, Twitter is likely to be regarded as a virtual 'tree hole' in that it serves to allow individuals to express the repressed feelings from their everyday lives. Users may therefore deliberately engage in or intensify conflictual Twitter interactions in order to present opinions and emotions that they are hesitant or reluctant to express in the real world because of the constraints imposed by social norms. Fifth, Twitter users appear to treat abrupt withdrawals as an approach to publicly display their control of interactions. By ending a conversation unilaterally, Twitter users interfere with their opponents' rights to speak in an unmarked manner; they also prevent their opponents from disregarding, provoking or disdaining them through the use of withdrawals, thus preserving their own face in advance.

Another feature that flouts face-to-face norms but is quite commonly observed on Twitter is that users regularly escalate conflict. This study showed that Twitter users tend to employ conflictual linguistic strategies, including disagreements and personal attacks, to object to their opponents' viewpoints or to humiliate their opponents or to combine both in certain contexts in which they feel that their views are not being heard or are being sidelined. The analysis demonstrates that disagreement strategies function in largely the same way in Twitter and in face-to-face contexts. In contrast, personal attacks on Twitter display a number of unique characteristics. First, posts that serve as a personal attack are commonly and frequently distributed in conflictual Twitter interactions across different contexts. The analysis shows that there are a number of posts in the corpus that include offensive and/or abusive terms, but the use of such bald-on-record attacks are generally avoided in offline contexts, even during face-to-face arguments. There are also many posts that are constructed by the speaker as an indirect attack and that do not involve any structural properties of personal attacks such as swear words. These mitigated attacks, which are commonly observed in face-to-face arguments, also make up an important percentage in my corpus and are however interpreted as a personal attack by other Twitter users. Second, personal attacks from the perspective of the recipient also play an important role in driving the development of conflictual Twitter interactions, which contrasts with the stereotypical assumption that attacks hinder communication. They function to maintain the continuation of conflictual interactions by largely guaranteeing people's continued engagement in the exchange and by expanding the reach of interactions through increasing the number of participants. They also serve to greatly intensify the conflict by escalating interpersonal disputes into an antagonism between opinion camps or even social groups and to change participants' strategic goals from arguing a topic to struggling for the right to speak. These findings, to some extent, also indicate that personal attacks appear to serve as an acceptable strategy on Twitter. Their occurrence seems to be largely dependent on the discursive context and it is difficult to identify them based on fixed linguistic features as used in current computational detection models.

Apart from using disagreements and/or personal attacks in a bald-on-record manner, Twitter users also employ linguistic politeness approaches to modify their opposition, including pseudo-politeness, 'off-record' politeness, negative politeness and positive politeness. By using pseudo-politeness, they publicly announce their provocative or sarcastic

intents to emphasise their opposition, thus greatly intensifying the conflict. In order to oppose their opponents without taking any responsibility, Twitter users tend to employ 'off-record' politeness to disguise their intents to damage face. This is done by providing a clue by which the recipient may interpret the speaker's linguistic behaviour as being polite. Negative politeness and positive politeness are also employed by Twitter users to minimise the damage to their opponents' face in conflictual interactions. However, the protection of their opponents' negative face appears to be merely a by-product of the users' primary purposes of preserving their own images. In contrast, positive politeness seems to be only used when Twitter users sincerely seek to terminate the conflict with their opponents. The reason is that it requires the users to strike a temporary or even permanent compromise with their opponents and to abandon their dominance of opinions, such as giving up persuading the opponent to accept an opposing point of view and allowing different viewpoints equally to exist in the interaction. As a result, positive politeness is found to be effective in making the recipient feel that their face has been adequately preserved in a conflictual Twitter interaction but also to be quite rare on Twitter. This contrasts with the cases of 'off-record' politeness and negative politeness. Although these two approaches are generally considered as useful for preserving the recipient's face in face-to-face arguments (e.g., Kakavá, 1993; Muntigl & Turnbull, 1998; Rees-Miller, 2000; Locher, 2004; Ben-Menachem & Livnat, 2018), in most cases, they fail to serve this function on Twitter and even tend to slightly escalate the conflict. This suggests that the threshold for what constitutes courtesy has become much higher on Twitter than in face-to-face settings, and people have to exert greater effort when attempting to express truly linguistically (positively or negatively) polite behaviour.

The analysis finally revealed that Twitter users are cooperative in using linguistic strategies to do conflict. This contrasts with the stereotypical assumption about interactions in pragmatic research which maintains that inharmonious interactions inhibit the cooperation between participants (e.g., Brown & Levinson, 1987/1978; Leech, 2014/1983). This cooperation is accomplished by fulfilling four tasks: first, Twitter users assume the role of the recipient to interpret the (previous) speaker's turn design, including identifying what types of strategies are used by the (previous) speaker and speculating about what functions the speaker intends their turn to serve. Second, they have to formulate their own strategic goals, including deciding whether or not to respond to the (previous) speaker and whether

or not their response goes along with their interpretation. Third, they act as the (next) speaker, constructing their response on the basis of their interpretation and their own communicative intent. Finally, they publish their response in the next turn and leave it for other Twitter users to interpret freely. In this way, by constantly switching from being a recipient to being a speaker, Twitter users cooperate with each other and thus progress the turn-taking in conflictual interactions. This indicates that the conversational mechanism used in DA to explain face-to-face interactions still pertains to the case of conflictual Twitter interactions.

This study further shows that the turn constructions of both the speaker and the recipient play an equally important role in this cooperation because mismatches commonly occur between their designs. Sometimes, the recipient does not seem to comprehend or interpret the (previous) speaker's use of strategies and/or strategic goals in the same way as the speaker; even if the recipient interprets a post in the same way as the (previous) speaker, they may also deliberately avoid constructing the next turn in accordance with that interpretation to pursue specific interactional goals. This mismatch assigns the speaker's and the recipient's turns with different implications when analysing the role that a turn serves in an interaction. The nature of each turn is a complex outcome of the interpretation of the previous turn and their own strategic goals. Apart from this, the speaker's turn design also reveals their personal expectations of how their turn functions in the subsequent interaction; the recipient's turn design indicates what role the speaker's turn actually serves in the interaction. This suggests that in order to reveal how a turn progresses an interaction, it is useful to investigate how its recipient designs a response based on the (previous) speaker's turn rather than solely examining the makeup of the speaker's original turn in terms of its use of linguistic politeness and types of conflictual interactions.

8.2 Implication of the study

Taking into account the above-mentioned findings, this section discusses the contributions that the present study makes from four angles: (1) the development of pragmatic research (see Section 8.2.1), (2) language use and communication (see Section 8.2.2), (3) analytical frameworks of CMC (see Section 8.2.3) and (4) practical implications (see Section 8.2.4).

8.2.1 Conflict vs. pragmatic research

Pragmatic research on interactions particularly focuses on harmonious interactions and pays less attention to conflictual ones. Researchers tend to have a default notion built on Grice's Cooperative Principles (1975), asserting that people's cooperation is dictated by their agreements and friendliness with each other (Sifianou, 2019). They are also likely to value social harmony highly in interactions, especially in English discourse, as a result of the Anglocentric nature of much of their research and their reliance on English-based data (Klein, 2009; Wierzbicka, 2011; Levisen, 2016, 2019). Impoliteness studies, a subfield of pragmatics researching uninhibited linguistic behaviour, have grown rapidly over the past decade. However, they have not had a significant impact on shifting the excessive emphasis placed on harmonious linguistic behaviour in pragmatics research to date. Researchers are generally concerned with the consequences and intentions of actions that are associated with impoliteness. Pragmatic research on conflict pays a much greater attention to various types of uninhibited linguistic behaviour that serve to constitute a social process known as conflict, considering not only the impacts and intents of these actions, but also their internal nature and, more importantly, the language logic embedded within these actions in interactions. As a result, conflictual interactions are still typically treated as chaotic and being of little significance in pragmatic research on human interactions. However, this study challenges this stereotype and shows that conflictual interactions are equally important to study as harmonious interactions.

The analysis revealed that people's turn designs in conflictual interactions can be deconstructed into various types of linguistic strategies and modifiers. According to their strategic goals, people select suitable conflictual linguistic strategies to communicate the primary messages about opposing their opponents; they then combine these strategies with appropriate linguistic politeness modifiers to convey some relational content such as their mood, virtual identities and relationships to the surroundings. Since people appear to be cooperative in conflictual interactions, these linguistic actions that they perform to engage in conflict are also characterised by regular patterns. According to the discussion in Chapter 5 (see Section 5.2), people have to take turns from being a recipient of the previous turn to being a speaker of the next turn. It is through this successive role shift that conflictual interactions produce consistent turn-takings, thus maintaining their continuation and development. The analysis also found that the need for preserving the face that has been

damaged by the previous turn serves as an intrinsic cause for this role switch. All of these findings indicate that, as in harmonious interactions, people's language use in conflictual interactions is also orderly heterogeneity because it has distinct characteristics and regular compositions and performs relatively stable functions. This discovery challenges the commonly held assumption of much of pragmatic research, which typically adopts a "structural-functionalist social paradigm" (Bell et al., 2016, p. 401).

Functionalists view "the order and stability… as essential for the maintenance of the social system" (Mesthrie, 2009/2000, p. 28) and believe that this social harmony is preserved by social norms. They argue that "unless norms are shared, members of society would be unlikely to cooperate and work together" (Mesthrie, 2009/2000, p. 29). This leads to the belief that "conflict is just a temporary disturbance of the social order" (Mesthrie, 2009/2000, p. 30), which has been the implicit assumption in much of pragmatic research. However, this study suggests that conflictual interactions not only characterise CMC but also possess an orderly heterogeneity. This finding aligns with the Marxist viewpoint in sociology, arguing that "conflict is a common and persistent feature of society" (Mesthrie, 2009/2000, p. 30) and leads to the need for a reassessment of the essence of interactions. Contrary to the functionalist belief that each social group benefits from a smooth operation of society, Marxism emphasises the "fundamental difference of interest among social groups" (Mesthrie, 2009/2000, p. 29) and provides the foundation for the field of critical discourse analysis (CDA) established by Norman Fairclough. CDA focuses on "the way social power abuse, dominance and inequality are enacted, reproduced and resisted by text and talk in the social and political context" (Van Dijk, 2003, p. 352), aiming to reveal the relationship between language structures and social structures. They treat the imbalance of social power between individuals as an intrinsic factor that drives people's linguistic practices, which was also found in this study.

It appears from the analysis that information about whether people are objecting to their opponents' viewpoints or attacking the opponents is the primary message that they want to convey in conflictual interactions (see Section 7.1 in Chapter 7). However, the analysis also reveals that people's fundamental purpose for doing conflict is not to exchange information. When investigating cases of personal attacks and linguistic politeness, it is common to find that Twitter users engage in a tug-of-war with their opponents to assert their power on a given topic. In order to protect their right to speak and announce their

dominance of opinions, Twitter users interfere with their opponents' freedom of speech, which in turn motivates these opponents to violate their previous speakers' rights out of a desire to defend their own rights. This study, therefore, argues that human communication is a linguistic practice where interlocutors perform actions against their own discursive power and that of the other party. In harmonious interactions, in order to maintain a friendly and peaceful social relationship, people acknowledge the other party's right to speak by appropriately conceding their own right. In conflictual interactions, however, individuals tend to interfere with others' freedom of speech, with the aim of protecting their own right to speak. This interactional perspective presents another possible way beyond CDA for integrating the Marxist viewpoint into pragmatic research, which may facilitate a stronger connection between pragmatic research and social problems.

8.2.2 Technology vs. human communication

Linguists have noted, as early as the late 20th century, that CMC lacks many of the informational cues that are normally present in face-to-face interactions, including gestures, postures, facial expressions, intonation, social context information, social feedback and personal information (e.g., clothing, social status, occupation, age, gender, name) (e.g., Daft & Lengel, 1983; Baron, 1984; Kiesler et al., 1985; Siegel et al., 1986; Myers, 1987; Dubrovsky et al., 1990; Reid, 1991; Aycock, 1995; Parks & Floyd, 1996; Postmes et al., 1999). This has led to a long-running debate among linguists regarding the impact of technology on language change (e.g., Reid, 1991; Sproull & Kiesler, 1991; Baym, 1995; Garton & Wellman 1995; Aycock 1995; Parks & Floyd, 1996; Postmes et al., 1999; Murthy, 2013; Haugh & Sinkeviciute, 2019; Graham, 2019; Sifianou, 2019). Based on the findings in conflictual Twitter interactions, this section therefore discusses in detail whether and how technology changes human language use and communication.

Discussion 1: technology does not change…

Technology does not reduce the expressiveness and complexity of natural language. The informational cues that were considered as absent in virtual interactions (e.g., gesture, posture, facial expression, intonation, social context information, social feedback, personal details) do not disappear in CMC. They are just manifested in a different manner from that found in face-to-face interactions. For example, Chapter 7 (see Section 7.2.5) found that in order to communicate vivid and subtle emotions or feelings that are unavailable from texts,

such as friendliness, happiness, mockery and sarcasm, Twitter users mimic different types of laughter in face-to-face interactions such as smiles and laughs or sneers and bitter smiles. This is achieved by employing various textual or multimodal features such as emoji (e.g., 🙂, 😆, 😂), emoticons (e.g., ':)'), texts that describe the sound of laughing or the state of smile (e.g., *hhhhhh, lol*) and images or GIFs that depict smiling faces and body movements.

Although Twitter algorithms enable and protect the concealment and alteration of identities, Twitter users often treat revealing identities as an effective approach to taking sides and increasing the persuasiveness of their opposing positions. For example, in the conversation between Users A19 and C19 in Excerpts 55–58 (see Section 6.2 in Chapter 6), User A19 repeatedly reveals their difficulty in wearing a mask as an autistic patient to refuse User C19's insistence regarding mandatory mask wearing. Another case in point are the arguments between two gender groups triggered by @Jessica's tweet in Excerpt 105 (see Section 7.2.1 in Chapter 7). @Jessica sets up a confrontation between the two gender groups in her tweet, posing the question: "What would you do if all men had a 9pm curfew?" and reveals her identity as a feminist by addressing men and women separately and by adding hashtags such as "feminism" and "#manprivilege". Although most participants in the above conversation did not explicitly state their gender identities, as @Jessica did, their language use is able to clearly communicate which gender group they support and, in many cases, even reveals the identities they intend to display on Twitter. These virtual identities may not be the same as users' real-life identities, but they are used by the users to construct another ego that exists on the Internet and to project a social status and lifestyle that users desire, but cannot achieve, in real life. The analysis also shows that Twitter users are not isolated because social feedback and social connections are not absent from Twitter as people stereotypically assume. By using the function of following, they participate in and observe different virtual communities that they are interested in. They also have multiple ways to access and disseminate information. They can obtain the public's reactions and opinions towards a certain issue by browsing posts and find resonance with individuals who share a similar viewpoint or interest. They can also convey their thoughts and beliefs and display their lifestyle and philosophy by tweeting and retweeting posts and develop interactions with others by using the reply function.

In addition, DA argues that in face-to-face settings, after the current speaker has published the turn of their own, the interaction moves towards a TRP where participants

need to re-allocate the role of the speaker. The party who assumes the role of the (next) speaker, subsequently, is required to respond to the previous speaker and to issue the next turn, thus accomplishing a turn-taking. The analysis of conflictual Twitter interactions (see Section 5.2.1 in Chapter 5) indicates that this conversational mechanism has not been altered fundamentally by Twitter algorithms because interlocutors on Twitter still need to take turns being a speaker and to issue a new turn. It is therefore reasonable to assume that this mechanism, as maintained by DA, is also applicable to interpersonal interactions that occur on other online platforms. However, the specific manifestations of some key elements of interactions such as turns, turn-taking and the method of selecting the next speaker, appear to have been changed by Twitter algorithms, which leads to the following discussion of how technology has altered human communication.

Discussion 2: technology has changed…

This study found that technology has contributed to changes in people's linguistic practices in various aspects, including the form and scope of natural language, the structure of human communication and people's language ideologies. First of all, technology promotes the evolution of natural language, making it more vivid and diverse. It is impossible to deny that in the early stage of CMC, people's language use was relatively monotonous compared to that in face-to-face interactions because of technological constraints such as word limits and a limited number of signs. For example, there were few options for people to communicate their emotions, facial expressions, gestures and postures in the early phase of CMC, the most notable of which was the emoticon. Emoticons are "punctuation marks, letters and numbers used to create pictorial icons that generally display an emotion or sentiment" (Grannan, n.d.), such as ':)', 'T^T' and '^0^'.[32] Emoji, the first pictograph of faces and bodies, were created in the 1990s in Japan and became globally popular in the 2010s due to the publication of Unicode 6.0 and its application on the iPhone (Novak et al., 2015; Emoji research team, 2015; Miller-Ott et al., 2016; Zhao, 2018). Since then, a variety of multimodal features such as images, GIFs and videos, have been developed and widely used in CMC. This suggests that even in cases where interactions are restricted, people are still able to create conditions for expression and interaction using the tools that are currently available, thus creating new types of linguistic variables. In response to

[32] Accessed, 22nd February 2022: https://www.britannica.com/story/whats-the-difference-between-emoji-and-emoticons

people's need to communicate, new technologies are developed by web companies and computer scientists, which in turn encourages people to create new linguistic variables and thus promote language change and variation.

Second, technology alters the patterns and structure of human interactions, but this change is stongly influenced and shaped by the algorithms of different online platforms. Twitter algorithms, for example, enable human interactions to be displayed, recorded and saved as text. According to DA, a turn in face-to-face conversations is a stretch of words uttered by one speaker from start to finish, which means that its occurrence and termination are able to be signalled aurally by the (dis)appearance of the speaker's voice. However, a turn on Twitter refers to all messages published by a user in response to the previous speaker (see Section 4.2.1 in Chapter 4). This results in a situation where the speaker allocates their words to two posts (due to their strategic goals or the word limit of a tweet) without specifying that their words are continuing. In this case, the recipient is very likely to be misled regarding the scope of the speaker's turn. It is even more challenging for the recipient to realise how the (previous) speaker issues their second reply because the speaker may comment on their own first reply or respond twice to their previous speaker (see Section 4.2.1 in Chapter 4).

The written nature also results in an indefinite scope of Twitter interactions, both in terms of their duration and the number of participants and posts (see Section 4.2.1 in Chapter 4). Users can temporarily set aside a Twitter interaction and return to it whenever they are available, without having to schedule a large time block as they do in face-to-face interactions. Twitter interactions thus usually last for a couple of days, which are generally longer than those in face-to-face contexts. A case in point is the argument between Users B2 and C2 from Excerpts 26 to 38 (see Section 5.3.2 in Chapter 5) on whether prayer is suitable for victims of Hurricane Dorian. A day later, as shown in Excerpt 126 (see Section 7.3 in Chapter 7), User E2 objects to User C2's challenge of the existence of God. They accuse User C2 of being cold-blooded and lacking in compassion, "May Allah strengthen your mind and soften your heart" (Lines 100–101), which is then endorsed by User B2 at the next level, thus reigniting the debate between Users B2 and C2 on the second day.

On Twitter, a single post usually triggers multiple replies and these replies further elicit additional replies (see Section 4.2.1 in Chapter 4). By doing so, interpersonal

interactions on Twitter are shown as a tree-like structure (Scheffler, 2017), consisting of multiple one-on-one and one-to-many interactions, which challenges the rule of a successful face-to-face conversation, that of "one party talking at a time" (Sacks, 2004, p. 37). Twitter algorithms also collapse lengthy interactions on the first page of a tweet. To see the full conversation and participate in a lengthy interaction, Twitter users have to click the reply button multiple times (see Section 6.4.2 in Chapter 6), which results in the majority of Twitter interactions being short, containing only two turns. Therefore, broadcasting messages on Twitter is much easier than establishing a lengthy one-on-one conversation with fixed interlocutors as in the case of face-to-face contexts.

Third, the changes in language use and interactional patterns also further lead to the emergence of new language ideologies. Twitter algorithms establish a virtual world where people are free to create a new persona by hiding their real-life identities and to interact with strangers in a different manner than in face-to-face interactions. Twitter also enables users to interact via mobile devices at scattered times, which makes the actions of entering and exiting Twitter serve as a switch between the real and the virtual worlds. Twitter users are forced to connect the action of clicking on the Twitter logo to access or leave the site with the action of switching between their real and virtual personas. This tends to create a sense of disconnect from reality, leaving people with a misconception that once they enter the Twitter website their behaviour will no longer be constrained by social norms, administrative regulations or laws. In order to protect individuals' freedom of expression, Twitter's own guidelines for regulating and punishing its users for their uninhibited linguistic behaviour are likewise quite lax, which further weakens the imposition of social norms on users' language practices. Twitter has thus been treated as a space where individuals can express opinions, attitudes and emotions free from social constraints. They do not need to worry about receiving social moral condemnation, ruined reputations or administrative penalties for speaking rudely, aggressively or harmfully, such as using swear words, eliciting social hatred or spreading rumours, as in face-to-face interactions. Therefore, people on Twitter appear to use bald-on-record attacks and pseudo-politeness approaches more frequently and boldly than in face-to-face conversations.

Twitter algorithms also prioritise posts that contain bald-on-record attacks and/or pseudo-politeness in post ranking because these posts tend to receive a large number of comments. Despite the low frequency of these language strategies on Twitter, they can still

be noticed by a large number of Twitter users. People are used to receiving harmful language aurally in face-to-face interactions. On Twitter, however, they are forced to access such information visually by using a new sense organ, the eyes. These few bald-on-record attacks and pseudo-politeness thus have a profound impact on people's perceptions of Twitter, forming the stereotype that Twitter is rude, significantly enhancing the threshold of what can be interpreted as polite on Twitter and decreasing the threshold of what language is acceptable on Twitter.

Since conversations are based on turn-taking, it is very common for participants to expect a reply from the other party. However, unlike WhatsApp and Messenger, Twitter does not indicate whether or not a message has been checked if this message is not responded to. It is also impossible for Twitter users to supervise or urge their recipients to respond by using eye contact, facial expressions or gestures, as in face-to-face conversations. People on Twitter have to lower their expectations of receiving a response and instead to output their personal opinions one-sidedly; otherwise, they need to exert pressure on their recipients to require them to respond. For example, the speaker may raise an explicit question and wait for an answer from the recipient. They may also damage the recipient's face in a way that cannot be ignored, forcing the recipient to respond in order to preserve their own face, which is quite common in the corpus. However, the recipient tends to have the same concern, resulting in a vicious circle. The tension between interlocutors increases as their argument continues. Even if one party attempts to terminate the argument by opening a closing sequence, the interaction would eventually revert to that vicious cycle. This is because the closing procedure that is taken as courteous in face-to-face conversations overtly challenges the default of a conflictual interaction, namely, opposition. The recipient has to draw a reference to testify their assumption that the (previous) speaker is not innocent. Therefore, on Twitter, an abrupt withdrawal become the most effective and convenient way to end and/or leave an argument.

8.2.3 Analytical frameworks of CMC

A direct consequence of the lack of attention paid to technology as an important influencing factor in CMC contexts is that CMC studies primarily employ the theoretical frameworks and methodological approaches that were designed for face-to-face interactions without adaptations. To address this issue, this study suggests that CMC studies must be

carried out using a solid framework that crucially takes into account the technological affordances of the platform and the forms of communication that they engender. First of all, it is necessary to gain a thorough understanding of the structure of virtual interactions on the platforms under study, particularly their types, in-built features and how these features implement different forms of interactions. Insights into these matters are best based on the literature that is available on these platforms and through conducting pilot studies. Second, researchers should examine the benefits and limitations of different linguistic theories used to analyse face-to-face interactions and adapt them based on the characteristics of the online interaction under study. For example, this study suggests that the combination of Brown and Levinson's (1987/1978) facework and DA constitutes an appropriate approach for analysing the makeup and structure of conflictual Twitter interactions. However, it also noted that the current frameworks of these two theories are limited. When applying facework to the analysis of linguistic politeness, Brown and Levinson posited that the discursive function of language use is consistent with its literal function, which is however challenged by pseudo-politeness on Twitter. This study therefore abandoned Brown and Levinson's (1987/1978) notion of politeness and redefined it based on the corpus as a linguistic modifier used to mitigate or intensify FTAs (see Section 7.1 in Chapter 7). In terms of DA, the analysis originally found that APs are appropriate for investigating how the SPP is formed based on the FPP, including how it responds to the content of the FPP and how they match each other structurally. However, this approach is unable to provide insights into how the SPP actually functions in interactions. This is quite crucial in investigating how an interaction forms and develops because, as shown in this study, there are often mismatches between the designs of two adjacent turns. The role that the speaker expects their turn to play is often at odds with what it actually serves as in the interaction. The study thus questioned the use of APs as a means for analysing interactions and suggests that three-turn units provide a more suitable approach (see Section 5.2 in Chapter 5).

Adopting a multidisciplinary approach in CMC research is also needed. Researchers should borrow useful theories and methods from multidisciplinary fields to address the challenges posed by network technologies to enrich classic theories and approaches used for studying face-to-face data. For example, since this study took Twitter algorithms as a social factor that influences Twitter users' language use and communication, it conducted a sociolinguistics-oriented pragmatic analysis. It also integrated corpus linguistics into this

framework because it provides a suitable framework for researching online data of high representativeness. Specifically, this study used a corpus linguistics approach to collect interactional data on Twitter and to build the Conflictual Twitter Interaction Corpus. It also conducted questionnaire surveys to obtain information that is not easily obtained from interactions to supplement the interactional data in the corpus, such as Twitter users' interpretations, attitudes and opinions. To analyse the corpus, this study combined a qualitative analysis by using DA and Brown and Levinson's (1987/1978) facework with a simple quantitative analysis by using frequency proportion and Pearson correlation coefficient. The aim is to assist researchers to gain an overview of the data in the corpus and to quickly identify the issues that require a more in-depth qualitative analysis. In addition, researchers are also supposed to carefully assess the theories and methods from multidisciplinary fields that have been selected, considering what parts of these theories and methods need to be modified or adapted in order to better address the research question. For example, existing theories in corpus linguistics, which are more appropriate for collecting and analysing large volumes of isolated tweets that are used for a quantitative analysis, are not fully applicable to this study that aims to establish a corpus of conflictual Twitter interactions used for a qualitative analysis. To address this issue, this study reformulated the meaning of 'representative' and created a reasonable classification criterion for tweets based on research findings in sociolinguistics and natural language processing on language communities and data classification. It then established a step-by-step data collection procedure that specifies the type of data to be collected such as textual data, metadata and multimodal data, and a procedure for how the data is to be collected and archived. It also developed a self-designed annotation standard and verified the validity of this standard through questionnaire surveys.

8.2.4 Detection and regulation of uninhibited linguistic behaviour

To identify and moderate uninhibited linguistic behaviour on their platforms, social media, including Twitter and Facebook, currently rely on a combination of user reports and employee reviews. They generally set a standard of language use, particularly in relation to language that poses a serious threat to cyber security and social stability, such as hate speech, and encourage users to report posts that fail to meet the standard. This seems to indicate that social media providers are concerned with their users' interpretations. However, after receiving a report, they allocate the report to several employees who are

responsible for content moderation. These employees then review the report based on their own experiences, generate a decision based on the majority rule and proceed to take appropriate actions based on the decision such as deleting the post. In other words, decisions are made solely according to the interpretations of social media employees, regardless of whether or not the employees' interpretations are consistent with those of users. This leads to the level of verbal hygiene of social media being highly influenced by employees' personal competence, which includes issues such as their skills in language analysis, their personal understanding of the standard and their workload and efficiency. It appears that applying computational language detection models that are developed specifically for identifying online hate speech to content moderation tasks may provide a solution because they are able to significantly improve the efficiency and accuracy of hate detection. However, to my knowledge, the vast majority of social media companies have not yet publicly stated whether they use such models to identify hate speech and other uninhibited linguistic behaviour.

According to the current literature, these models are used primarily for optimising and developing natural language processing techniques and are less effective in dealing with uninhibited linguistic behaviour in a more general sense (e.g., Nobata et al., 2016; Davidson et al., 2017; Saleem et al., 2017; Mondal et al., 2017; Salminen et al., 2018ab; Watanabe et al., 2018; Salminen et al., 2020). The present study argues that this issue is due to two main factors. First, the analysis shows that hate speech only serves as a special form of personal attack that is employed to invocate social hatred against certain ethnic or vulnerable groups. There are, however, much broader causes of conflictual interactions, other than social hatred that is deeply rooted in sociohistorical collisions. They may be triggered by any type of topic, even non-offensive and non-sensitive ones, such as CONVERSATIONs 22 and 25 in the corpus which deal with complaints about Apple products (see Excerpts 72–73 and Excerpts 83–86 in Section 6.4 in Chapter 6). To engage in a conflict, in addition to hate speech, people also use personal attacks in a more general sense and disagreements that are simply used to refute the opponent's point of view. Second, hate speech is generally considered as a set of fixed linguistic features with identical functions and its detection is primarily based on textual features such as keywords, BOW, N-grams and POS. This contrasts with the findings of Chapter 5 (see Section 5.2) and Chapter 6 that the function of language use varies depending on interactional contexts and does not necessarily correspond to its

apparent linguistic features. The analysis shows that the function of a turn depends entirely on the recipient's interpretation and their own strategic objectives, but the recipient's response and actions often do not project the use of conflictual linguistic strategies and linguistic politeness in that turn from a structural perspective.

Due to these limitations, current computational detection models are more appropriate for dealing with conflictual interactions that stem from hatred and discrimination with deep sociohistorical roots and also involve abusive and offensive terms such as taboo language and swear words. They are far from being able to distinguish conflictual exchanges from others in a more general sense, much less being used as a predictive tool for the intensification and outbreak of an online conflict. This requires detectors to interpret the literal meaning of a post in the absence of obvious linguistic features and to deduce the communicative intents of this post from contextual and discursive information. It also requires the computer to predict the discursive function of a post and the conflict trend and to identify the possible turning points of conflict that signify the transition from the stage of exchanging information and building relationships to the stage of arguing points of view, then to the stage of mutual aggression and humiliation. In addition, this study found that communication on Twitter is not primarily intended to maintain harmonious relationships and/or to facilitate mutual understanding; rather, seeking information and announcing opinions are of greater importance. It was noticed that, although personal attacks generally serve to severely exacerbate conflict and threaten social stability like disagreements, they also play an important role for CMC users in expressing their opinions and presence. Personal attacks, which were treated as detrimental to communication in the stereotype, are instead found to be a major strategy used by people to register their anger and attitudes, to announce their dominance of opinions, to seize the right to speak and to drive the development of conflictual interactions. This study argues that personal attacks are being perceived as an acceptable linguistic strategy by CMC users in virtual contexts. This means that if personal attacks are identified and managed indiscriminately, they are likely to trigger greater social conflict about freedom of expression. Determining the timing of controlling the use of personal attacks therefore becomes another difficulty. At that moment, participants and other CMC users generally do not perceive the third-party control as an interference in their freedom of speech but begin

to be willing to request a third-party regulation. To address all of these issues, there is still a need for cross-disciplinary cooperation, which is discussed in the next section.

8.3 Constraints and outlook

This section discusses constraints of the current study and identifies directions for future research relating to conflictual interactions on the Internet. Since well-designed software fails to collect interactive data, to differentiate conflictual exchanges from interactions or to qualify language use, communicative intent and interactional function of posts in conflictual interactions, this study had to perform all of these tasks manually. This results in a relatively small corpus. This study is thus unable to compensate for the lack of a specific assertion in previous research regarding whether conflictual interactions are more prevalent on Twitter by measuring their frequency and distribution but merely concluded based on a qualitative analysis that conflict does characterise Twitter interactions. This issue, for now, cannot be addressed by making the classification approach used to annotate the corpus in this study suitable for quantification, since this approach relies primarily on a comprehensive understanding of the contextual information and discursive functions of posts, which is beyond what current natural language processing techniques can handle. In addition, although this study has established a reasonable data classification criterion to ensure representativeness, the corpus is highly likely to exclude some unique language uses or interactional patterns that are typical of particular contexts. To resolve these issues, linguists need to collaborate with computer scientists in order to optimise the natural language processing techniques used for qualitative analysis, as well as to improve the efficiency of data collection, processing and analysis by applying these techniques to their interactional research.

This study is also constrained by the lack of comparative examples of arguments in face-to-face settings. Although the analysis also discusses the differences between Twitter interactions and face-to-face interactions, this comparison was based on previous studies. They mainly examined the use of disagreement strategies in face-to-face settings and linguistic politeness in harmonious interactions and interactions that aim to build a harmonious relationship in face-to-face settings. Although many disagreement studies also investigated the use of linguistic politeness in conflictual interactions, there is still a lack of research on the use of linguistic politeness in conflictual interactions in a more general

sense. There are even fewer studies in linguistics that address the use of personal attacks in face-to-face contexts due to the difficulty in collection. Therefore, this study is unable to compare the frequency, the language use and interactive patterns of verbal conflict in virtual and face-to-face environments in more detail. To address this issue, linguists first have to tackle the challenge of gathering oral conflict data, such as improving the techniques used for linguistic observation or surveys, and collecting alternative data such as that obtained from experimental research and talk shows. They then need to establish comparative corpora of face-to-face and online interactions on different platforms and to further divide the corpora into sub-corpora of conflictual and harmonious exchanges in different settings, respectively and to conduct a comparative study based on these corpora in more depth.

Lastly, this study revealed the linguistic mechanisms of conflictual interactions on Twitter by qualifying interactional data and demonstrated that these mechanisms can also be applied to other CMC platforms from a theoretical perspective. However, it does not further substantiate this argument by using interactional data from other social media such as YouTube, Facebook and TikTok, nor does it delve into what adaptations should be made when applying these mechanisms to the research of other online interactions. This will be an important area for further investigation based on the findings of this study. This study also does not provide any immediate solutions about how to detect the emergence of conflict, how to predict the course of a conflictual interaction and the point at which it escalates from an opinion argument to an antagonism, or how to mitigate, control or resolve a conflict. However, the research in this thesis provides a solid basis for examining these issues. Further investigation requires not only more qualitative research into language use across multiple online platforms but also the cooperation of researchers from different disciplines. For instance, linguists may work with computer scientists to optimise the techniques of natural language processing regarding interactional data and contextual information, to explore parameters that are effective for identifying exchanges with different intensity and developmental tendency in a conflictual interaction, to examine the relationship between various environmental factors (e.g., duration, the number of participants, the range of diffusion) and the development of conflict. Linguists may also collaborate with cognitive scientists to conduct experimental research to explore issues such as what factors influence people's interpretation of posts, what factors cause people to abandon social norms but to blatantly provoke or attack opponents, and what stages of conflict management are most

effective in avoiding both the recurrence of the current conflict and the emergence of new social conflict.

Appendix 1 Interactions

Interaction 1

Line	User Account	Text	Level	Turn
01	@James	[1R]I can't even afford food and rich people are eating gold on theirs. Fuck capitalism and fuck the rich [2R][QUOTE: https://twitter.com/InsiderFood/status/1145157506869710848]	tweet	1
02				
03				
04				
05	User A1	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
06				
07				
08	User B1	[1R]you're 15, anti-ACAB, a centrist, and a furry. Your opinion is completely and utterly invalid	2nd	5
09				
Omit Lines 10–20				

Interaction 2

Line	User Account	Text	Level	Turn
01	@John	1R]A desperate cry for help[Crying face][Crying face][Crying face][Crying face]#HurricaneDorian #Abaco #Bahamas Lord please help us [2R][VIDEO: https://twitter.com/i/status/1168227588533293056]	Tweet	1
02				
03				
04				
05	User A2	[1R]Father in Heaven, please safeguard Your people in the path of Hurricane Dorian. Loose Your mighty angels to minister unto each of them at this time. Fill them with Your peace that passes all understanding. Keep them together in love. In Jesus' Name, Amen.	1st	2
06				
07				
08				
09				
10	User B2	[1R]And the animals and wildlife too, Lord, created by you.[Folded hands][Folded hands][Folded hands]	2nd	3
11				
12	User C2	[1R]I'm assuming that whilst everyone is praying to their god to save them, no-one is considering that an all powerful god would have prevented the storm in the first place. Doesn't sound like a very compassionate deity to me.	3rd	4
13				
14				
15				
16	User B2	[1R]Relax Atheist, your day of reckoning will come… everyone's does. But it's funny how, when one's life or a loved one's life is in danger people turn to asking their God to save them. But that would never be you, right??? You'll have to wait and see…	4th	5_1
17				
18				
19				
20	User C2	[1R]You are quite correct – it is funny how people to turn to a god to save them from bad things whilst never wondering why their loving all-powerful god allows the things to happen in the first place. The hypocrisy baffles me.	5th	6_1
21				
22				
23				
Omit Lines 24–30				
31	User C2	[1R]Presumably they'll use the all encompassing get out clause that their god moves in mysterious ways…. hmmm….	5th	6_2
32				
33	User C2	[1R]Oh, and by the way – obviously I hope that those that have been affected by the storm are ok – because I'm not a horrible bellend…. I just don't see that praying is likely to help them in any way whatsoever. I was brought up with religion and it has done nothing for me.	5th	6_3
34				
35				
36				
37				
38	User B2	[1R]Then why troll the comments of those that do believe? ... have you nothing better to do, that can help people that are afraid and in danger of dying in a hurricane, who are asking us to pray for them?[1WS] This is your solution?	6th	9
39				
40				
41				

(To be continued)

Line	User Account	Text	Level	Turn
42 43 44 45 46	User C2	[1R]I've not trolled anyone. I've not said anything insulting about any individuals. Just making the point that I fail to see how it will actually help anyone. They are better supported by actual action rather than prayers - whether it be financial or structural aid. Just my opinion.	7th	10_1
47 48 49 50	User B2	[1R]Sweetie, these people in the videos where the sorm is now can't get help until it's over. THEY are asking for prayers bc they are scared.[1WS] Do this for them, it's called COMPASSION, and it won't hurt you!!![Folded hands]	8th	11
51 52 53 54 55	User C2	[1R]I can assure you I'm thinking of them, & I'm hoping that they are ok. I just won't pray to a god I don't believe in. I apologise if I caused you offence - I can assure you it wasn't intended that way. I just keep seeing posts about praying, yet bad stuff still happens in droves.	9th	12
56 57 58 59 60	User B2	[1R]Unfortunately the world is made up of good and bad things. We are also given free will to chose a given path bc God didn't make us to be mindless stick figures to experience life in a bubble.[1WS] We make mistakes, and sometimes we need help, or we're thankful.	10th	13_1
61 62 63 64 65	User B2	[1R]... so we give thanks to him. It's a comforting way to live in faith. Especially as we realise we're not in control of ANYTHING!Someone once said, even if there's a chance there is no God, living in faith is a nice way to live this life we're given- It's true	10th	13_2
66 67 68	User C2	[1R]I'd happily donate - and often do - to support those in peril or danger. That surely benefits them more than me telling them I'm thinking of them???	7th	10_2
Omit Lines 69–70				
71 72	User C2	[1R].... and to be honest - you threatened me with a day of reckoning!	7th	10_3
73 74 75	User B2	[1R]No. I didn't "threaten" you with a day of reckoning... it's just a fact of life, it happens to everyone one day. You will learn what your honest reactions are when it happens to you..	8th	15
76	User D2	[1R]They aren't asking you to pray for them. FFS, piss off.	7th	16_1
77	User D2	[1R]Your prayers help no one.	7th	16_2
78 79 80 81 82	User C2	[1R]Everyone has their right to an opinion - believers and non-believers. If you believe, then that is obviously fine. I just don't understand why people look to something to save them, when if their beliefs are correct, that god should prevent the bad stuff in the 1st place.	5th	6_4
83 84	User D2	[1R]I would never ask for any God's help if I was in danger as god does not exist. Bloody bible-thumping weirdo.	5th	17
Omit Lines 85–96				
97 98 99	User B2	[1R]And whilst you sit there "assuming", what exactly did you do for the girl in the video who's PLEADING FOR PRAYERS???? If you don't want to pray for her... THEN DON'Tperiod[Neutral face]	4th	5_2
100 101	User E2	[1R]May Allah strengthen your mind and soften your heart. Ameen	4th	22
102 103 104	User B2	[1R]That's all I was trying to tell him![1WS] "Soften your heart", was well put, brother[Smiling face with smiling eyes][Folded hands][Red heart]	5th	23

(To be continued)

Line	User Account	Text	Level	Turn
105 106 107 108 109	User C2	[1R]Now look here folks - just because I don't believe in a religion doesn't mean that I have no compassion. I firmly believe that you can be a good person by doing good deeds - and I do. It is arrogant & insulting to infer that I am hard hearted just because I don't have a god!	6th	24
110 111 112 113 114	User B2	[1R]That's NOT what anyone was saying.[1WS] I suggest you work on your reading comprehension, or perhaps you're just crying wolf, and not as dumb as you make out?[1WS] Either way, stop attacking people for their beliefs.... it's weak and pathetic. Ciao[Smiling face with open mouth and smiling eyes]	7th	25
Omit Lines 115–123				

Interaction 3

Line	User Account	Text	Level	Turn
01 02 03 04 05 06 07	@Brian	[1R]Australia. This is the road between Canberra and the Coast I think yesterday. The idea that it is somehow not right 'to talk about #ClimateChange at this difficult time' is frankly convenient bulllshit. Click on image to see what life is like now for the #firefighters[1WS]. [2R][PIC: https://twitter.com/TwoPaddocks/status/1208207260477882369/photo/1]	tweet	1
Omit Lines 08–27				
28	User C3	[1R]Why thin skinned?	9th	10
29 30	User D3	[1R]Because you think being called a loser is abuse[Face with tears of joy]	10th	11_1
31 32 33	User D3	[1R]It doesn't change the fact that your one example is a misquote that got fixed to mean the opposite of what you wanted it to mean	11th	11_2

Interaction 4

Line	User Account	Text	Level	Turn
01 02 03 04	@Independent	[1R]Alexandria Ocasio-Cortez grills former Exxon scientists on oil giant's climate change denial [2R][VIDEO: https://twitter.com/i/status/1187719206562910209]	tweet	1
Omit Lines 05–08				
09 10 11 12	User B4	[1R]that's odd. to say it clearly, #ExxonKnew[1WS], no, industry is NOT able to do what nature is doing, that is the problem. Nature regenerates and creates, industry destroys. end of cake. we are not God. [2R:ER] [3R]2	2nd	3
13 14 15 16 17	User C4	[1R]I'm not sure I follow the point [2R:ER] [3R]1) #ExxonKnew at the very least that there was a very high risk of bad things happening, but they pushed the narrative that there was none [4R]2) Bad things have started happening [5R]3) It's not their fault at all because who could know for sure?	3rd	4

(To be continued)

Line	User Account	Text	Level	Turn
18	User B4	[1R]Sorry, but perhaps you listen to the interrogation of the scientist who worked for #exxonknew , if you have any questions , that they knew exactly what a)they were doing b) was comming c) that they are guilty. [2R]https://twitter.com/YEARSofLIVING/status/118748872668 15072256 [3R][Quote: https://twitter.com/Independent/status/11877192065629102 09]	4th	5
19				
20				
21				
22				
23				
24				
25				

Omit Lines 26–27

Interaction 5

Line	User Account	Text	Level	Turn
01	@Brian	[1R]Australia. This is the road between Canberra and the Coast I think yesterday. The idea that it is somehow not right 'to talk about #ClimateChange at this difficult time' is frankly convenient bulllshit. Click on image to see what life is like now for the #firefighters[1WS]. [2R][PIC: https://twitter.com/TwoPaddocks/status/12082072604778823 69/photo/1]	tweet	1
02				
03				
04				
05				
06				
07				
08	User A5	[1R]Arson	1st	2

Interaction 6

Line	User Account	Text	Level	Turn
01	@Mary	[1R]The intolerable & intrusive sound of #Islamic imperialism in Brooklyn, New York. [2R:ER] [3R]I thought the USA was a #Secular nation... [4R][VIDEO: https://twitter.com/i/status/908092802487132160]	Tweet	1
02				
03				
04				
05	User A6	[1R]Church Bells are a thing of the past due to atheist complaints. You can bet they don't care if muslims disturb the peace.	1st	2
06				
07				

Omit Lines 08–11

Interaction 7

Line	User Account	Text	Level	Turn
01	@James	[1R]I can't even afford food and rich people are eating gold on theirs. Fuck capitalism and fuck the rich [2R][QUOTE: https://twitter.com/InsiderFood/status/114515750686971084 8]	tweet	1
02				
03				
04				
05	User A7	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
06				
07				
08	User B7	[1R]Oh so a furry want to talk[Loudly crying face][Loudly crying face][Loudly crying face][Loudly crying face]	2nd	3
09				
10	User A7	[1R]Please don't pin this on furries. Everyone is susceptible to propaganda and it's completely unrelated to the fandom.	3rd	4
11				

Interaction 8

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				

(To be continued)

Line	User Account	Text	Level	Turn
Omit Lines 04–08				
09 10 11 12 13	User A8	[1R]A seatbelt saves your life, a mask isn't going to do that genius. There's no legitimate proof that a mask makes any difference. It restricts your breathing though, because I've had to wear one on public transport, but don't expect us to be thrilled about it.	3rd	3
Omit Lines 14–27				
28 29	User D8	[1R]Someone: *agrees with someone [2R]Tyrone: "sToP gEtTiNg InVolVeD"	9th	9
30 31 32	User A8	[1R]You sound weak bro. You don't need to get involved, it's a two way convo, and you're getting in the way. I'm sure your Mum taught you manners, use them.	10th	10
33 34 35	User D8	[1R]I'm sure your mum didn't teach you manners because you'd rather people die from a virus than swallow your pride... [2R]Twitter is a free platform[Thumbs up]	11th	11
Omit Lines 36–41				

Interaction 9

Line	User Account	Text	Level	Turn
01 02 03	@Susan	[1R]Even in a disaster no one wants the vegan food. [2R][PIC: https://twitter.com/hillisthekillis/status/903720326231719936/photo/1]	tweet	1
04 05 06	User A9	[1R]It's ironic because if everyone were eating this food, the natural distaster may not have even happened[Woman shrugging]	1st	2

Interaction 10

Line	User Account	Text	Level	Turn
01 02 03 04	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
05 06 07	User A10	[1R]I would protest it – because 99% of men[ren't the animals you are trying to convince everyone they are. Stop being sexist. #twowaystreet	1st	2
Omit Lines 08–43				

Interaction 11

Line	User Account	Text	Level	Turn
01 02	@PinkNews	[1R]Makeup has no gender[Person tipping hand] [2R][VIDEO: https://twitter.com/i/status/1287036885953650688]	tweet	1
03	User A11	[1R]awww so cute!	1st	2
04 05	User B11	[1R]Yeah, if you think pedophilia and child abuse is cute [2R:ER] [3R]In which case, I'd call the FBI on ya	2nd	3

Interaction 12

Line	User Account	Text	Level	Turn
01	@Paul	[1R]amazon rainforest: how can i help? need advice on how i can actually make a change here; these pictures are breaking my heart... this is one of the most important ecosystems on earth. is there a fundraiser, a call to action, anything i can leverage my audience for on this? [2R][PIC: https://twitter.com/LoganPaul/status/1164064145198600193/photo/1]	tweet	1
02				
03				
04				
05				
06				
07				
Omit Lines 08–22				
23	User B12	[1R]Why are you assuming he would... Visit the forest. I'm not going to argue with you, we all need to do our part, no matter how shitty of a person we are. Famous people SHOULD be putting their money/celebrity towards fighting this. His fan base will follow his lead	4th	5_1
24				
25				
26				
27				
28	User C12	[1R]Because it's a youtuber known for taking things to the extreme, just like his brother (though in different ways).	5th	6
29				
Omit Lines 30–49				
50	User B12	[1R]What more can he do? He is literally asking what he can do in the tweet you responded to. If this were anyone else you would even notice but because you dislike LP you want to hinder helping the planet. Can't wait for global warming to kill us all	5th	5_2
51				
52				
53				

Interaction 13

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A13	[1R]But you knew you'd have to wear one when u took the job so shut up	1st	2
05				
06	User B13	[1R]She can see into the future? She knew when she took the job there was gonna be a pandemic? Make it make sense lmao	2nd	3
07				
08				
09	User A13	[1R]No doctors and nurses have always worn masks bro	3rd	4
10	User B13	[1R]Not constantly all day on a ward they haven't[Rolling on the floor laughing]	4th	5
11				

Interaction 14

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A14	[1R]I found the Latin for 'do not be a dick' so much more appealing. [2R][PIC: https://twitter.com/IamPaulGibson/status/1287042602143973376/photo/1]	1st	2
05				
06				
07				

Interaction 15

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				

(To be continued)

Line	User Account	Text	Level	Turn
04 05 06 07	User A15	[1R]If I can wear a mask, you can, blah…... Blah…..blah…...blah…...blah…... [2R]Yawn[Sleeping face][Sleeping face]. Every man and his dog say the same thing on here[Face with tears of joy].	1st	2

Interaction 16

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
04	User A16	[1R]It's your job	1st	2
05 06	User B16	[1R]and it's your job as a decent human being to wear one too :)	2nd	3
07	User C16	[1R]Nope	3rd	4_1
08 09	User C16	[1R]If anything not wearing 1 is as healthy because people stay well clear :)	3rd	4_2
10	User B16	[1R]that's... not how it works	4th	5
11 12 13	User C16	[1R]If people are happy to wear em, let em, if people are happy not wearing em, let em. not all shops are following the "guidance" and neither should they	5th	6
14 15	User B16	[1R]it's been advised by MEDICAL EXPERTS for a reason. they work to protect people, being a selfish twat doesn't	6th	7

Interaction 17

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
04 05 06 07	User A17	[1R]Could you please explain to me why when I was in hospital in May for three days most of he doctors and nurses didn't wear masks?[2R:ER] [3R]I'm guessing you won't.	1st	2
08	User B17	[1R]You were at a sh*t hospital	2nd	3
09	User A17	[1R]You people are insane. Gullible, unthinking sheep.	3rd	4
10 11 12 13 14	User B17	[1R]Do your own research on RELIABLE websites like the cdc or any other site with ".org" or ".edu" in the url. Don't rely on ".com" sites like you clearly have been. There's times to stand out from the crowd and there's times to save lives by following others	4th	5
15 16 17	User A17	[1R]I do.[1WS] I presume you can read? [2R][PIC: https://twitter.com/Thisaintright12/status/1287109704007680002/photo/1]	5th	6
18 19	User B17	[1R]You're giving me information on patients dying in the UK?	6th	7
20 21	User A17	[1R]And you're telling me to check cdc?![2R:ER] [3R]Works both ways[Face with rolling eyes]	7th	8
22 23 24	User B17	[1R]I'm telling you that masks work and you just sent me information on patients dying in the uk?? I'm done talking about this bot[Face with tears of joy]	8th	9

Interaction 18

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A18	[1R]It was ok to not wear one Thursday but now all of sudden Covid has got worse overnight[Thinking face]You do not wear a mask for 13 hours and if you want to believe magically that wearing a mask 4 months into a pandemic is gona help now then thats yout choice but don't critical of others.	1st	2
05				
06				
07				
08				
09	User B18	[1R]'Don't be critical of others' I'm definitely going to be critical of people who won't wear a mask for a short amount of time in a shop (unless they are exempt), why is it so hard for some people to do I don't understand	2nd	3
10				
11				
12				
Omit Lines 13–30				
31	User B18	[1R]Why are you so against wearing a mask for 20 mins tops???	8th	9
32				
33	User A18	[1R]Cos it wasn't relevant 4 months ago and isn't now	9th	10_1
34	User C18	[1R]I bet you clapped on Thursdays for the NHS doing a great job, just to disrespect them online and potentially cause a second lockdown. You can call ppl who wear masks sheep's all you want, but I think the word for someone who won't wear them in public is selfishtwat.	10th	11
35				
36				
37				
38				
39	User D18	[1R]aww fuck up sean and just wear a mask you absolute nonce. As more and more people are now out in public and coming into contact with each other then we need to try reduce the risk of spreading it and a mask does that! Don't want to wear a mask? Then don't go out! Simple!	10th	12
40				
41				
42				
43				
44	User E18	[1R]Shut up you fucking moron	10th	13_1
Omit Lines 45–49				

Interaction 19

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A19	[1R]some people can't wear them though, "if i can where one, so can you" isn't exactly true	1st	2
05				
06	User B19	[1R]who	2nd	3
07	User A19	[1R]people with asthma and other breathing problems, autistic people, some people with anxiety etc... obviously if you can where one, you should but some people don't have to because they genuinely can't handle it	3rd	4
08				
09				
10				
11	User C19	[1R]I know a ton of people with severe asthma, as well as nurses with asthma, masks don't stop you from breathing	4th	5
12				
13	User A19	[1R]they don't stop you but can make you feel like it's harder to breathe, i also know people with asthma that try to avoid going to shops due to being unable to cope wearing one	5th	6
14				
15				
16				
17	User C19	[1R]I mean, if you really are struggling that much, don't go to the shop. I have asthma and the last thing I'd want is a respiratory infection[Woman shrugging]	6th	7
18				
19				

(To be continued)

Line	User Account	Text	Level	Turn
20	User A19	[1R]that's what we're doing, we keep trying to get deliveries instead but they're that busy atm so sometimes we don't have a choice but to go...	7th	8
21				
22				
23	User C19	[1R]but then you should still put on a mask..	8th	9
24	User A19	[1R]i'm autistic and i can't handle it for more than 5 minutes so i cant... i'd much rather wear a mask than risk getting covid but it not that easy..	9th	10
25				
26				
27	User C19	[1R]stop being selfish and stay at home then xoxo cba to argue with people that only care about themselves[Face with rolling eyes]	10th	11
28				
29				

Interaction 20

Line	User Account	Text	Level	Turn
01	@washingtonpost	[1R]Breaking news: Airstrike at Baghdad airport kills Iran's most revered military leader, Qasem Soleimani, Iraqi state television reports [2R][QUOTE: https://www.washingtonpost.com/world/national-security/defense-secretary-says-iran-and-its-proxies-may-be-planning-fresh-attacks-on-us-personnel-in-iraq/2020/01/02/53b63f00-2d89-11ea-bcb3-ac6482c4a92f_story.html?utm_campaign=wp_main&utm_medium=social&utm_source=twitter]	tweet	1
02				
03				
04				
05				
06				
07				
08				
09				
10	User A20	[1R]He was a legit Islamic terrorist.[2R:ER][3R]My God, you guys aren't even trying hide your hatred of America anymore.	1st	2
11				
12	User B20	[1R]He was not an "international terrorist". He was a high ranking military officer of a nation. That's not the same as OBL or Al-Baghdadi. This is like Iran taking out our Chairman of the Joint Chiefs. Do you know anything or just spout stupid MAGA taking points?	2nd	3
13				
14				
15				
16				
17	User C20	[1R]who did you blow to get a blue check mark?	3rd	4
18	User B20	[1R]Wait that's how you got yours? I thought it was my decades of military service and sparkling personality. Who did you have to do?	4th	5
19				
20				
21	User C20	[1R]storm Twitter headquarters in San Francisco.	5th	6
22	User B20	[1R]Again...was not aware of that as a technique. I did it when you still had to apply and prove you were a human.	6th	7
23				
24	User C20	[1R]me too hon	7th	8

Interaction 21

Line	User Account	Text	Level	Turn
01	@Paul	[1R]amazon rainforest: how can i help? need advice on how i can actually make a change here; these pictures are breaking my heart... this is one of the most important ecosystems on earth. is there a fundraiser, a call to action, anything i can leverage my audience for on this? [2R][PIC: https://twitter.com/LoganPaul/status/1164064145198600193/photo/1]	tweet	1
02				
03				
04				
05				
06				
07				
08	User A21	[1R]all i seen was logan paul and rainforest and my heart sank	1st	2_1
09	User A21	[1R]no hate btw[Folded hands][Black heart]	2nd	2_2
10	User B21	[1R]Mate stop criticising when even you dont bring anything to the table. So please just get a life.	2nd	3
11				
12	User C21	[1R]And you did?	3rd	4

Interaction 22

Line	User Account	Text	Level	Turn
01	@David	[1R]coming soon: #iphone 15 - It's All Cameras!™ #AppleEvent [2R][PIC: https://twitter.com/starboots_/status/1171490143942758401/photo/1]	tweet	1
02				
03				
04				
05	User A22	[1R]People think this is funny, but it seriously takes me awhile to get over seeing something like this.	1st	2
06				
07	User B22	[1R]are you a fucking ape?	2nd	3
08	User C22	[1R]so don't look at it[Clown face]	2nd	4
09	User D22	[1R]Imagine being so attention hungry u tell other people and tell urself ur scared of holes lol	2nd	5
10				
11	User E22	[1R]People aren't gonna feel sorry for you because you got in your feels over a post on twitter	2nd	6
12				
13	User F22	[1R]Don't be a fucking weakling then. Take control over your body	2nd	7
14				

Interaction 23

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
04	User A23	[1R]Oh I was also walked right through a 'red zone' accidentally with no PPE.[1WS] Guess what.. I'm still here. [2R:ER] [3R]STOP SPREADING FEAR.	1st	2
05				
06				
07	User B23	[1R]Shut it karen	2nd	3
08	User A23	[1R]You're hard.[Face with rolling eyes]	3rd	4
09	User B23	[1R]You're so solid that you won't wear a piece of cloth to help the cause but you be an idiot	4th	5
10				
11	User A23	[1R]'Help the cause'[Face with tears of joy] [2R:ER] [3R]Good one.[1WS] Can you spell potato? [4R][PIC: https://twitter.com/Thisaintright12/status/1287031639323881474/photo/1]	5th	6
12				
13				
14				
15	User C23	[1R]They still wont get it[Face with tears of joy]	6th	7
16	User B23	[1R]Your kids minging	7th	8
17	User D23	[1R]If only you were still a cum stain on your dads boxers pal	8th	9_1
18	User D23	[1R]Little billy shrimp dick verbally attacking a kid[Face with tears of joy]	9th	9_2
19				
20	User B23	[1R]Cry more [2R][GIF: https://twitter.com/i/status/1287064860434538499]	10th	10
21				
22	User A23	[1R]What a sad little keyboard warrior you are.[Face with tears of joy]	11th	11
23				
24	User B23	[1R][PIC: https://twitter.com/CoreyMFC/status/1287072402552303616/photo/1]	12th	12
25				
26				
27	User A23	[1R]Hilarious. [2R:ER] [3R]Have you thought about therapy?	13th	13
28	User B23	[1R]Says the woman who doesn't believe in a virus you mong	14th	14
29	User A23	[1R]You're not even a good troll.[Face with tears of joy] [2R:ER] [3R]Language skills of a 5 year old.	15th	15
30				
31	User B23	[1R]Says the one who thinks they're cool not wearing a mask divvy mate	16th	16
32				

Interaction 24

Line	User Account	Text	Level	Turn
01	@BBCBreaking	[1R]UK High Street pharmacy chain Boots cutting 4,000 jobs and closing 48 stores in wake of coronavirus pandemic [2R][QUOTE: https://www.bbc.com/news/business-53348519?at_campaign=64&at_custom2=twitter&at_medium=custom7&at_custom4=2764B6B6-C1D6-11EA-BC8A-8B8C96E8478F&at_custom3=%40BBCBreaking&at_custom1=%5Bpost+type%5D]	tweet	1
02				
03				
04				
05				
06				
07				
08	User A24	[1R]Is @BootsUK refunding all the furlough money claimed if they're cutting 4000 jobs? @RishiSunak @hmtreasury	1st	2
09				
10	User B24	[1R]Why would they? People who had left their jobs and been fired were eligible to furlough, it wasn't given with the requirement that there be a job there at the end of it.	2nd	3
11				
12				
13	User A24	[1R]It was a Job Retention Scheme not a Furlough Fraugh Scheme. @hmtreasury should recover all the taxpayers money spent on furlough claims by employers who never intended to retain those jobs! #Boots #johnlewis	3rd	4
14				
15				
16				
17	User C24	[1R]But the gov would have paid the same money in benefits instead,[1WS] it gave companies time to review their operations and not make hasty quick decisions	4th	5
18				
19				
20	User B24	[1R]The cost would have been less in benefits. [2R]Furlough was 80% up to 2.5k. The benefit cap is £1.9k max (couple with children in London).	5th	6
21				
22				
23	User C24	[1R]But you are looking at it on a single person basis instead of taking into account the thousands of jobs and businesses it might have actually saved, most businesses would have gone bust without it, trouble is there is no real way of knowing	6th	7
24				
25				
26				
27	User B24	[1R]No mate, you said they'd have paid the same in benefits which was incorrect. [2R]Most businesses would have benefited from rent and business rate holidays and could have declared temporary stops to retain employees while avoiding costs for wages. [3R]There were and are other options.	7th	8
28				
29				
30				
31				
32	User C24	[1R]Rent holidays?[Rolling on the floor laughing]sounds like you are in a different country from all the businesses i know	8th	9
33				
34	User B24	[1R]You don't know businesses which pay rent?	9th	10
35	User C24	[1R]You said rent holidays, i know lots of businesses that tried but failed to get rent holidays and without the furlough scheme would have laid off staff and possibly gone bust	10th	11
36				
37				
38	User B24	[1R]Yes, government mandated and backed rent holidays. I'm talking about other options the treasury could have considered.	11th	12
39				
40				

Interaction 25

Line	User Account	Text	Level	Turn
01	@Nancy	[1R]As a trillion dollar company, you'd think that #Apple could figure out a way to make better cords. [2R][PIC: https://twitter.com/NotDonnaBrazile/status/1025051419907379201/photo/1]	tweet	1
02				
03				
04				
Omit Lines 05–70				

(To be continued)

Line	User Account	Text	Level	Turn
71 72 73 74 75 76	User B25	[1R]That's another joke. I don't spend my day customizing little widgets on my phone. I don't spend that time moving icons around in a specific pattern or rooting the phone. I use it to use apps, read email, browse the internet, use the SSH terminal. iPhone does it just fine.	6th	12
77 78 79 80 81	User A25	[1R]Yea you have never used a pixel with these usual fanboy responses. I do not have to customize a thing on my day to day, but when something annoys me right I can change it. Not just accept it because that's how the operating system is made.	7th	13_1
82 83 84 85 86 87	User B25	[1R]Nope, pixel's software is not that special compared to any other android, which I have used. How the hell am I a fanboy if I use products from all companies? Nothing annoys me about iOS, what's there to annoy me when all I do is just open apps? What do you need to change?	8th	14
88 89 90 91	User A25	[1R]Fanboys always go to the oh to much customization as the excuse that Apple does what they need it to do. I same old song. I am ending this tirade because you are a fanboy and opinions are something you cannot change.	9th	15
92 93 94	User B25	[1R]Calling someone who owns multiple products from multiple companies just shows how fragile your opinions and arguments are instead of using facts. Good on you.	10th	16
95 96 97 98 99	User A25	[1R]Dude, you have an opinion you are looking at this from your point of view, no one elses. I am ending this conversation because its fricken Twitter, and I am not wasting my time with a fanboy, who tries to displace brand love for tech innovation.	11th	17
100 101 102 103 104	User B25	[1R]I've been giving you facts 80% of the time. And I'm not missing out by much because I'm talking to an idiot who doesn't realize so. Especially when you can only call me a fanboy. Same old same old "you're a fanboy so I win" bullshit.	12th	18
105 106 107 108 109 110	User A25	[1R]Win what? A Twitter tirade? Listen my wife has an iPhone 6 regular person. She doesn't want upgrade to an 8 or X because the change sucks. That's regular person not a tech persona. I have a friend loves Apple and fanboys too, but cannot explain to me why he needs 128gb phone.	13th	19
111 112 113	User A25	[1R]by default chrome OS has an ssh terminal, and hell even Windows lets you do that with gitbash since.. Windows 7 before they adopted side loading linux	7th	13_2
Omit Lines 114–117				
118 119 120 121 122 123	User B25	[1R]Almost none of this is an opinion. [2R:ER] [3R]Facts: Mac batteries can be replaced. [4R]CNET put a boulder on a Mac and submerged it in water to test its durability. [5R]Display on X is the best of all smartphones according to displaymate. [6R]X has stacked logic boards with highest density	10th	22

Interaction 26

Line	User Account	Text	Level	Turn
01	@Brian	[1R]Australia. This is the road between Canberra and the Coast I think yesterday. The idea that it is somehow not right 'to talk about #ClimateChange at this difficult time' is frankly convenient bulllshit. Click on image to see what life is like now for the #firefighters[1WS]. [2R][PIC: https://twitter.com/TwoPaddocks/status/1208207260477882369/photo/1]	tweet	1
02				
03				
04				
05				
06				
07				
Omit Lines 08–18				
19	User B26	[1R]Then you haven't seen enough. Our country had an ETS that visibly reduced our emissions. Current govt abolished it, emissions steadily risen since. Govt ignored business case for funding firefighting equipment for 11 mths, now not enough eqp to fight fires.	4th	5
20				
21				
22				
23				
24	User C26	[1R]We could reduce our fossil fuel emissions to zero and it wouldn't make a scrap of difference. Co2 is naturally occurring. It's everywhere and it's in everything. Rising populations are the biggest environmental issue. We are all part of the problem.	5th	6
25				
26				
27				
28				
29	User B26	[1R]That is utter bullshit. Doing nothing is the cowards way. There are plenty of solutions but the wilfully ignorant are barring the way. Don't be part of the problem.	6th	7
30				
31				
32	User C26	[1R]I don't have any children. I'm doing my part.	7th	8_1
33	User C26	[1R]I'm not ignorant and I do what I can. Please tell me your solutions. Anyone can point the finger.	7th	8_2
34				
35	User C26	[1R]Plenty of solutions[Thinking face]	7th	8_3
36	User C26	[1R]Nine people have liked your tweet. Not one has mentioned a solution. Typical. I'm guessing none of you take unnecessary holidays, drive big cars, use air conditioners, have multiple tvs[Face with rolling eyes]	7th	8_4
37				
38				
39				
40	User D26	[1R]Joe, you're part of the problem. We have solutions, think big and not at the individual level. [2R:ER] [3R]Solidarity User B26 [Green heart]	8th	9
41				
42				
43	User C26	[1R]Still waiting to here the solutions	9th	10
44	User B26	[1R]I debated whether or not to bother given you're not genuinely interested in the answer and are looking for a gotcha moment but here goes... [2R][QUOTE: https://www.climatecouncil.org.au/full-list-of-fire-and-emergency-chiefs-recommendations-to-federal-government/]	10th	11
45				
46				
47				
48				

Interaction 27

Line	User Account	Text	Level	Turn
01	@James	[1R]I can't even afford food and rich people are eating gold on theirs. Fuck capitalism and fuck the rich [2R][QUOTE: https://twitter.com/InsiderFood/status/1145157506869710848]	tweet	1
02				
03				
04				
05	User A27	[1R]So you want everyone to starve to death? I get how wasteful that is but seriously communism is only good on paper.	1st	2
06				
07				
08	User B27	[1R]WHHHHHHHHAT IN THE FRESH HELL AHAHHA	2nd	3
09	User C27	[1R]bro you're 15	2nd	4
10	User A27	[1R]Bro I have the right of speech.	3rd	5

(To be continued)

Line	User Account	Text	Level	Turn
11	User B27	[1R]You can't even consent to having a girlfriend legally. sit down	4th	6
12				
13	User D27	[1R]Ur a furry so no.	4th	7
14	User E27	[1R]Well that doesn't have much to do with it, but carry on.	5th	8
15	User B27	[1R]Okay amphibi	6th	9

Interaction 28

Line	User Account	Text	Level	Turn
01	@washingtonpost	[1R]Breaking news: Airstrike at Baghdad airport kills Iran's most revered military leader, Qasem Soleimani, Iraqi state television reports [2R][QUOTE: https://www.washingtonpost.com/world/national-security/defense-secretary-says-iran-and-its-proxies-may-be-planning-fresh-attacks-on-us-personnel-in-iraq/2020/01/02/53b63f00-2d89-11ea-bcb3-ac6482c4a92f_story.html?utm_campaign=wp_main&utm_medium=social&utm_source=twitter]	tweet	1
02				
03				
04				
05				
06				
07				
08				
09				
10	User A28	[1R]We're so sorry for your loss, WaPo. How was he related to you again?	1st	2
11				

Interaction 29

Line	User Account	Text	Level	Turn
01	@BBCBreaking	[1R]Structure of Notre-Dame saved from total destruction, Paris fire official says https://bbc.in/2UAMlVn	tweet	1
02				
03	User A29	[1R]I think I'd be inclined to wait for a structural engineer's report to be fair.	1st	2
04				

Interaction 30

Line	User Account	Text	Level	Turn
01	@BBCBreaking	[1R]Structure of Notre-Dame saved from total destruction, Paris fire official says https://bbc.in/2UAMlVn	tweet	1
02				
03	User A30	[1R]Considering the damage and the speed with which the fire developed, I have difficulty in believing this is purely accidental	1st	2
04				
05	User B30	[1R]Really? It's that hard to believe a really old mostly wooden interior with draperies, tapestries and wood pews could go up in flames that quickly?	2nd	3
06				
07				

Omit Lines 08–09

Interaction 31

Line	User Account	Text	Level	Turn
01	@Matthew	[1R]An estimated 100,000 people in Berlin on #ClimateStrike. Today is massive [2R][VIDEO: https://twitter.com/i/status/1175026670593597441]	tweet	1
02				
03				
04	User A31	[1R]Looks like a bunch of Zombies to me.	1st	2

Omit Lines 05–06

Interaction 32

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
05	User A32	[1R]You do realize the vast majority of dudes don't bother or harrass women right? Your post comes off as incredibly condescending and shallow. How about instead of generalizing you start encouraging women and tell them how to protect themselves if they would like the option?	1st	2
06				
07				
08				
09				
10	User B32	[1R]Dylan... You clearly missed the point.	2nd	3

Omit Line 10

Interaction 33

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
05	User A33	[1R]Long, calm, solo nighttime walks. That would be amazing.	1st	1
06	User B33	[1R]you can have a long solo walk ...put something on other then 90% nude -.- like how dear you blame men for this -.-	2nd	2
07				
08	User A33	[1R]You're an idiot and I'm blocking you.	3rd	3
09	User C33	[1R]and i'm blocking you. i like this game.	4th	4

Interaction 34

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
Omit Lines 05–228				
229	User E34	[1R]No one is fine with curfew. You aren't, I'm not. The problem is that you were asked to think and listen to women. We would do mostly simple stuff that you take for granted. That's a lesson, I think. Just listen.	20th	64
230				
231				
232				
233	User D34	[1R]I understand that women feel threatened and scared, I get that. But I think that is terribly wrong to start that debate with - lets curfew all men. That message just makes feminists seem like man haters and doesn't even allow us to debate the issue that is underneath all of this.	21st	65
234				
235				
236				
237				
238	User C34	[1R]ok honestly i get what youre saying but i think some really did take it too literal. i[Face with tears of joy]agree compltly with what youre saying but i also agree with what i am getting at..i have no idea if any of this makes sense but yh	22nd	66
239				
240				
241				
242	User D34	[1R]Sure, I get that. I get what you are saying. Glad that you heard my point	23rd	67
243				
244	User I34	[1R]She didn't propose a curfew on all men, its a hypothetical scenario- essentially she asked women what they would do differently if they knew men wouldn't be out on the streets after 9. The point being to illuminate the ways male presence harms or affects women's everyday lives.	24th	68
245				
246				
247				
248				

(To be continued)

Line	User Account	Text	Level	Turn
Omit Lines 249–251				
252 253 254	User D34	[1R]Listen, that is nice, but to me it speaks about misandry present in these kinds of discussions, the message seems to me - only if there are no men, women will finally be or feel safe.	25th	70
Omit Lines 255–258				
259 260 261 262	User I34	[1R]And btw, if you think that's the message, perhaps reflect on why that might be, and listen to what women are saying about how unsafe they feel,[1WS] instead off writing it off instantly in your head.	26th	72_1
263 264 265 266 267	User I34	[1R]So we're talking about the existential threat men pose to women every day, and the huge psychological impact that has on women, and the material effect it has on their lives, & your response is to ignore that 100% and say "who will think of the men?" You're the problem, son.	26th	72_2
268 269 270 271 272	User D34	[1R]No,the problem is you. You intentionally keep saying all men pose threat to women, but that is simply not reality.Not all husbands, brothers,men in general, attack women. We need to talk and see which men attack women and why.But if you keep saying all men,there wnt of a talk.	27th	73
Omit Lines 273–430				

Interaction 35

Line	User Account	Text	Level	Turn
01 02 03	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
Omit Lines 04–152				
153 154 155 156 157	User A35	[1R]You work in a care home, so I have no issue with you wearing one in there, you are dealing with the most vulnerable, so you can't afford to take risks. But this whole business of going mandatory for everyone is unnecessary. People with breathing difficulties are exempted.	9th	58_2
158 159 160 161 162	User F35	[1R]I completely understand if someone is exempt from wearing a mask I just hope they take every other precaution necessary to make sure they don't have it/spread it. Get tested, wash your hands, don't cough on anything and if you have any symptoms self isolate.	10th	61_1
163 164 165 166 167	User F35	[1R]If you can wear a mask though you certainly should. I haven't been wearing one in shops until now because I got tested at work and I've been negative this whole time. I don't know why they're saying it's mandatory now but I want to keep the people I work for safe so I will	10th	61_2
168 169 170 171	User A35	[1R]Respect, I'm glad we can have a conversation without insulting one another. Your work looking after the elderly often goes under appreciated, and my Grandma relies on people like yourself, so thanks for all the good work you've been doing.	11th	62
Omit Lines 171–334				

Interaction 36

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
Omit Lines 05–52				
53	User C36	[1R]Women get lighter sentences for the same crime. Women are more likely to get custody of the kids. When it comes down to rape, a female victim is more likely to be taken seriously than male victim. And also the Duluth Model.	17th	18
54				
55				
56				
57	User D36	[1R]What crimes do they get lighter sentences? And I'm especially curious about when/how women are more likely to be taken seriously than men?	18th	19
58				
59				
60	User C36	[1R]Basically for any crime. And they are more likely to be taken seriously as rape/domestic abuse victims than male victims are. Ever heard of the Duluth Model?	19th	20
61				
62				
Omit Lines 63–83				

Interaction 37

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
Omit Lines 05–17				
18	User D37	[1R]But that's our point. Us 99% hate that 1% too and want to stand with you against them, not be conflated with them.	4th	5
19				
Omit Lines 20–33				
34	User C37	[1R]No one is saying "all men do this". We are saying "MOST women have experienced this at the hands of *a* man, and it needs to stop. Right now. Whatever it takes, this has to end. We deserve to not live in fear." If you have ideas, awesome! If not, just listen. Hear us.	8th	9
35				
36				
37				
38				
39	User D37	[1R]And you'll get not argument from me on that. I want to help be part of the solution, not treated as part of the problem. I don't feel like we're too far apart on this.	9th	10
40				
41				
Omit Lines 42–58				

Interaction 38

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
05	User A38	[1R]Just a friendly reminder.... The majority of guys are not rapists. As great as all of this sounds, let's not forget that you're lumping us all into a nasty group that most of us want/have nothing to do with.	1st	2
06				
07				
08				
09	User B38	[1R]Just a friendly reminder...why don't you pipe down and read the replies? [2R:ER] [3R]Well done on completely missing the point of the post [4R][GIF: https://twitter.com/i/status/1047455628929773568]	2nd	3
10				
11				
12				

Line	User Account	Text	Level	Turn
Omit Lines 13–17				
18 19	User C38	[1R]Not lumping you “all” into one group. Knowing we’re usually smaller & weaker than the bad guys.	2nd	7
20 21 22	User D38	[1R]Your point? The language being used on this thread is too generalised, and I am disgusted to know that people may have thought they would be raped by walking past me	3rd	8
23 24 25 26	User E38	[1R]Not all men are rapists, no, but all men have the potential to be rapists in the mind of a woman walking alone in the dark. Hang back to a safe distance, or cross the road to overtake. Small beer for you but would make a big difference to her.	4th	9

Interaction 39

Line	User Account	Text	Level	Turn
01 02 03 04	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]”What would you do if all men had a 9pm curfew?” [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
Omit Lines 05–06				
07	User A39	[1R]Why so negative	3rd	4
08 09 10 11 12 13 14	@Jessica	[1R]Honestly because it’s 4:30 in the morning and I am fucking exhausted from blocking literally thousands of entitled dude bros on this thread. [2R:ER] [3R]This thread is not for you to speak. It is for you to listen. And if you’re not up for it, then we don’t need you here. [4R][PIC: https://twitter.com/DanielleMuscato/status/1047780160446103552/photo/1]	4th	5
15 16 17 18	User B39	[1R]So your a feminist who wants EQUAL rights yet are telling someone of the opposite sex he has NO RIGHT to comment on a PUBLIC platform?? You do realise not all men are creeps and attackers	5th	6
Omit Lines 19–26				

Interaction 40

Line	User Account	Text	Level	Turn
01 02 03 04	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]”What would you do if all men had a 9pm curfew?” [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
Omit Lines 05–08				
09 10 11 12 13	@Jessica	[1R]It’s not just about rape and murder. [2R:ER] [3R]It’s about utter freedom from creeps, Nice Guys™, dude bros, misogynists, & mansplainers. The feeling when you & your friends are in the club bathroom & you can say what you REALLY think, you can be YOU—but anywhere, & everywhere.	2nd	3
14 15 16	User B40	[1R]Your delusional if you think [1WS]criminals would follow a curfew, all you would do is get rid of witnesses lol crim would skyrocket[Face with tears of joy]	3rd	4
17	User A40	[1R]Think you may have missed the point here Adam..	4th	5

Interaction 41

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
05	User A41	[1R]As a guy I wouldn't follow said curfew.	1st	2
06	User B41	[1R]yeahhhhhhhh man, this shit isn't about you.	2nd	3

Interaction 42

Line	User Account	Text	Level	Turn
01	@Paul	[1R]amazon rainforest: how can i help? need advice on how i can actually make a change here; these pictures are breaking my heart... this is one of the most important ecosystems on earth. is there a fundraiser, a call to action, anything i can leverage my audience for on this? [2R][PIC: https://twitter.com/LoganPaul/status/1164064145198600193/photo/1]	tweet	1
02				
03				
04				
05				
06				
07				
08	User A42	[1R]film a dead body in it	1st	2
09	User B42	[1R]Haha. So funny. So original. How about you STFU	2nd	3
Omit Lines 10–12				

Interaction 43

Line	User Account	Text	Level	Turn
01	@Jessica	[1R]Ladies, a question for you: [2R:ER] [3R]"What would you do if all men had a 9pm curfew?" [4R:ER] [5R]Dudes: Read the replies and pay attention. [6R:ER] [7R]#metoo #Kavanaugh #Cosby #feminism #maleprivilege #privilege	tweet	1
02				
03				
04				
Omit Lines 05–34				
35	User D43	[1R]Do you want to be part of the solution or what? Or are you going to just come on threads like this and complain.	9th	9_2
36				
37	User C43	[1R]No I'll follow your lead and try to tell other people to talk somewhere else.[1WS] Thanks and good day	10th	10
38				
Omit Lines 39–40				

Interaction 44

Line	User Account	Text	Level	Turn
01	@Jennifer	[1R]if i can wear a mask for 13 hrs on a hot ward tomorrow you can wear a mask for 20 minutes doing your shopping :) don't be a dick x	tweet	1
02				
03				
Omit Lines 04–08				
09	User A44	[1R]Hello - you're really not understanding how masks work. If someone wore a mask, but comes in contact with someone that has Covid-19, they're likely to get infected. If someone that's wearing a mask has Covid-19 comes in contact with someone that hasn't got the virus, they're...	4th	4_2
10				
11				
12				
13				
Omit Lines 14–22				
23	User B44	[1R]NHS England figures from last week indicate that 13 people in the entire country died from the virus. Overall, the IFR is 0.04% and falling, 99.96% survive and 80% have no symptoms. I think your mask has suffocated your brain.	5th	9
24				
25				
26				
Omit Lines 27–46				

Appendix 2 Questionnaires

Questionnaire 1. Annotation

Welcome to the survey! Hello, guys! I'm Xixiang, a Ph.D. student from Linguistics in UCD. You're being asked to participate in my Ph.D. research about how people do conflictual interactions on Twitter. The purpose of this questionnaire is to justify the rules of annotating the Conflictual Twitter Interaction Corpus. I would appreciate if you could spare a few minutes to answer the following questions carefully. Thank you for your support and cooperation!

You will be asked to complete a questionnaire. In the questionnaire, you will read 4 excerpts of the conversations taken from Twitter. You will be asked to read each conversation and to give your opinion about what you think is being communicated in each one. I would also like to really emphasise here that some of the language in the conversations may be seen as offensive. The survey should take you around 15 minutes to complete. Your participation in this survey is voluntary. You can opt out of the questionnaire at any time, for any reason and without any prejudice. Please be assured that your personal information will not be collected. Your answers will be kept completely confidential and only the statistical results will be shown in the thesis. If you have any questions about the survey, please email xixiang.zhao@ucdconnect.ie. By clicking the button below, you acknowledge that your participation in this study is voluntary and you are aware that you may choose to terminate your participation in the survey at any time and for any reason.

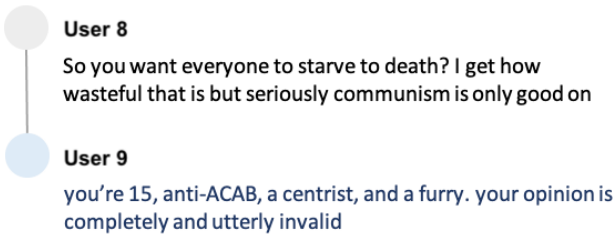
Please indicate your consent before continuing:

☐ I'm over 18 years old. I consent and start the survey

☐ I do not consent and wish to opt out of the survey

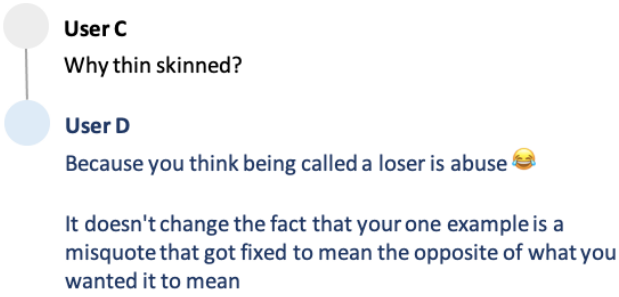
Section 1*: This part consists of three questions. Please briefly describe what the users marked in blue are doing. You can consider from the perspective of what the author of the message is doing with/to/for someone else (e.g., the author of the message is joking/negotiating a solution/sharing information with someone else; the author of the message is making a suggestion/request to someone else).*

1.1 What do you think User B is communicating to User A?



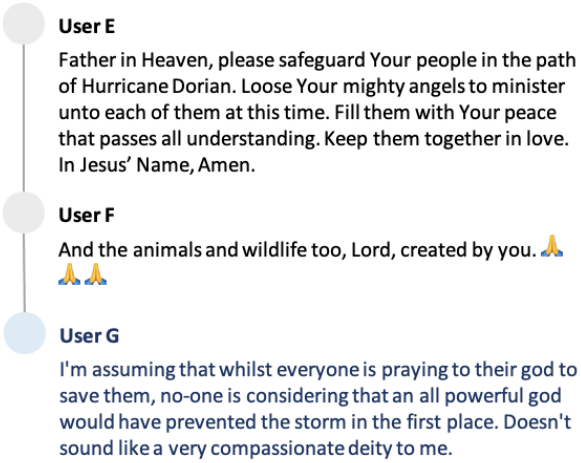
Answer: _____

1.2 What do you think User D is communicating to User C?



Answer: _____

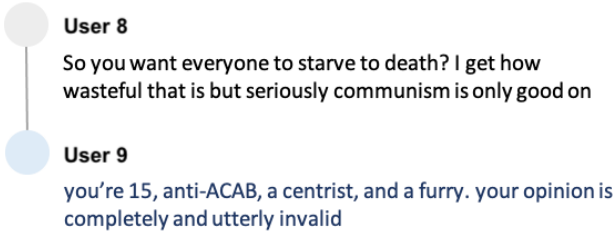
1.3 What do you think User G is communicating to Users F and E?



Answer: _____

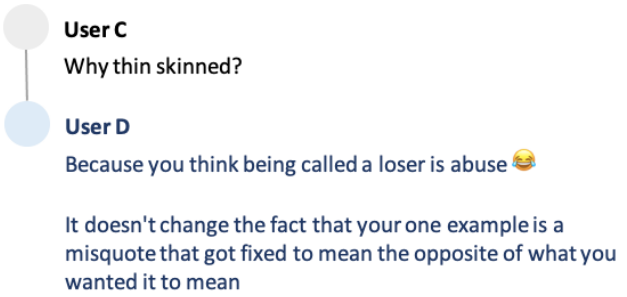
***Section 2**: This part consists of three questions. Please select the level of offence caused to you by the words of the users marked in blue. The level of offence ranges from 1 to 5. 1 is the least offensive and 5 is the most offensive.*

2.1 How offensive do you think User B's reply is to you?



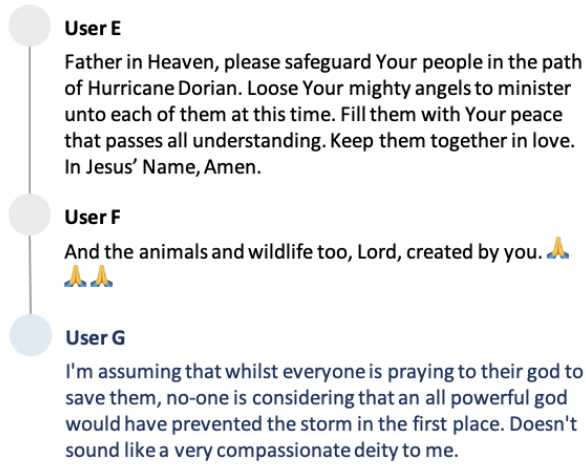
	1	2	3	4	5	
Least offensive	○	○	○	○	○	Most offensive

2.2 How offensive do you think User D's reply is to you?



	1	2	3	4	5	
Least offensive	○	○	○	○	○	Most offensive

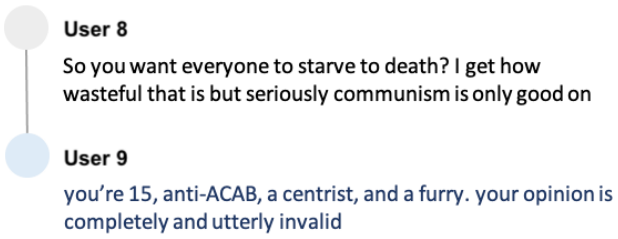
2.3 How offensive do you think User G's reply is to you?



	1	2	3	4	5	
Least offensive	○	○	○	○	○	Most offensive

***Section 3**: This parts contains three questions. Please annotate the action(s) performed by the users marked in blue to oppose someone else. You may choose one or more options. If you select 'Other', please detail your answers.*

3.1 What action(s) does User B perform against User A?

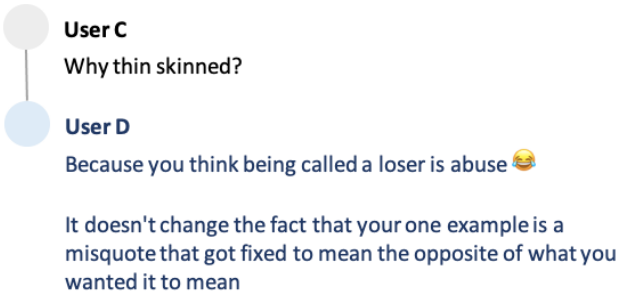


☐ Disagree

☐ Attack

☐ Other

3.2 What action(s) does User D perform against User C?

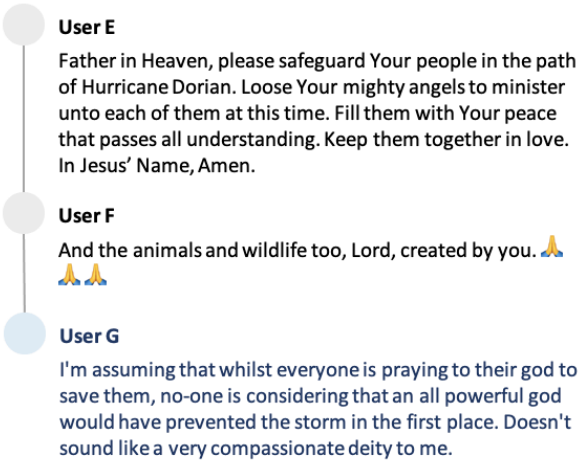


☐ Disagree

☐ Attack

☐ Other

3.3 What action(s) does User G perform against Users E and F?



☐ Disagree

☐ Attack

☐ Other

***Section 4**: In this part, you need to learn the annotation rules given. These will be used for the tasks in the next section.*

4.1 Please consider the annotation rules below: In the next task, you will be shown a tweet and will be asked to define what you think is being communicated from a list of 3 options:

(1) Disagree—this is where the author of the message disagrees with someone else's point of view.

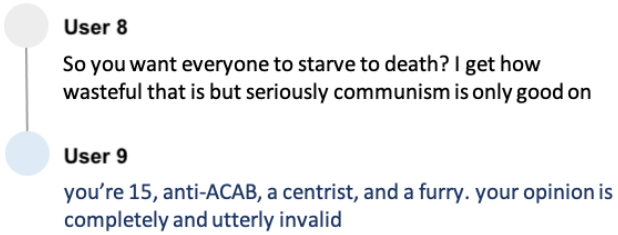
(2) Attack—this is where the author of the message attacks someone else with their tweet.

(3) Other—if the author of the message is doing something that is not a disagree or attack.

☐ I understand and continue to the next section

***Section 5**: This part contains three questions. You are asked, by using the rules you have learned, to annotate the action(s) performed by the users marked in blue. You may choose one or more options. If you select 'Other', please detail your answers.*

5.1 Please annotate the action(s) performed by User B.

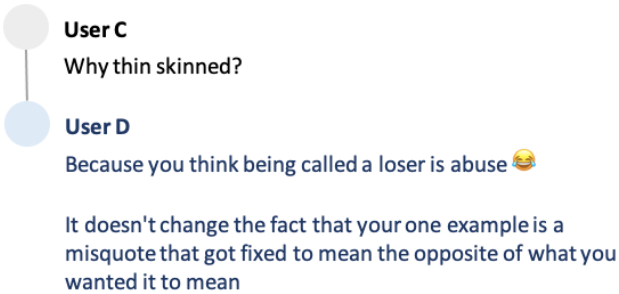


☐ Disagree

☐ Attack

☐ Other

5.2 Please annotate the action(s) performed by User D.

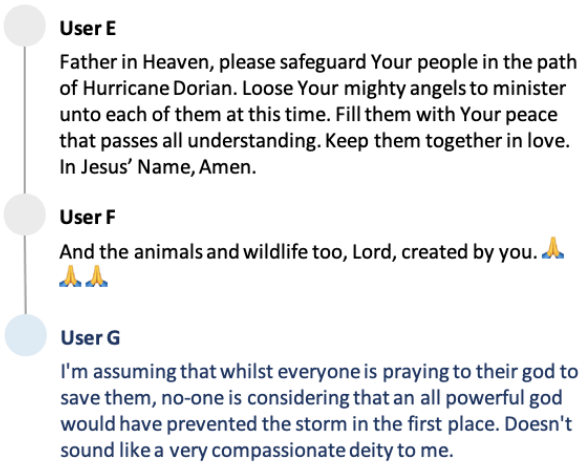


☐ Disagree

☐ Attack

☐ Other

5.3 Please annotate the action(s) performed by User G.



☐ Disagree

☐ Attack

☐ Other

Questionnaire 2. Attitudes

Welcome to the survey! Hello, guys! I'm Xixiang, a Ph.D. student from Linguistics in UCD. You're being asked to participate in my Ph.D. research about how people do conflictual interactions on Twitter. The purpose of this questionnaire is to gain insight into the reasons why Twitter users employ linguistic politeness when doing conflictual interactions. I would appreciate if you could spare 10 minutes to answer the following questions carefully. Thank you for your support and cooperation!

You will be asked to complete a questionnaire. This questionnaire consists of 4 tasks. You will first be asked to express your attitude towards the conversations on Twitter. Then, you will be asked to read a conversation taken from Twitter and to respond to the last user in that conversation. Next, you will be asked to follow my instructions to explain how you designed your reply. Finally, you will be asked to recall an experience of language use in conversations both on Twitter and in face-to-face. Please note that some of the language in the conversations and following questions may be appear to be offensive. The survey should take you around 10 minutes to complete. Your participation in this survey is voluntary. You can opt out of the questionnaire at any time, for any reason and without any prejudice. Please be assured that your personal information will not be collected. Your answers will be kept completely confidential and only the statistical results will be shown in the thesis. If you have any questions about the survey, please email xixiang.zhao@ucdconnect.ie. By clicking the button below, you acknowledge that your participation in this study is voluntary and you are aware that you may choose to terminate your participation in the survey at any time and for any reason.

Please indicate your consent before continuing:

☐ I'm over 18 years old. I consent and start the survey

☐ I do not consent and wish to opt out of the survey

***Section 1**: This task includes two questions. Please compare your experience of communicating with others in face-to-face contexts and on Twitter, and then briefly evaluate Twitter conversations.*

1.1 What is your favourite aspect of using Twitter?

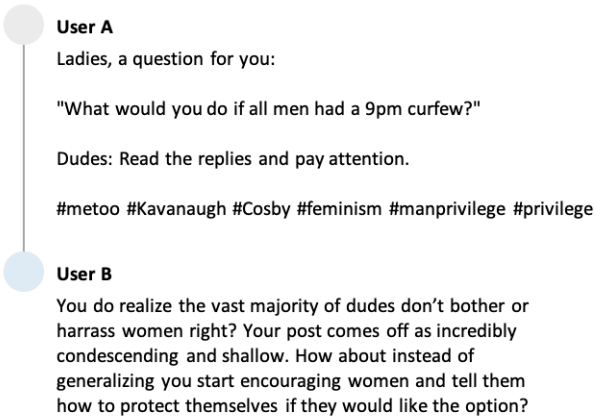
Answer: _____

1.2 What is your LEAST favourite aspect of using Twitter?

Answer: _____

Section 2*: This task only includes one question. Please assume that you are browsing Twitter, suddenly notice the conversation I provide below, and then respond to User B. You can support User B, or oppose User B, or attack User B. And you may also use dirty words or mention taboo topics. Please use the language format that you typically use on Twitter. Please do not exceed 280 words in your response.*

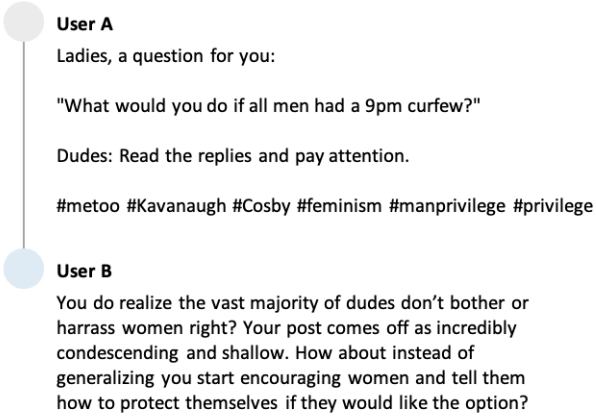
2.1 Please respond to User B.



Answer: _____

Section 3*: This task contains six questions. Please recall how you designed your reply in the previous task and follow the instructions to select the best option for each question. If you select 'Other', please briefly explain your answers.*

3.1 What did you think of User B's reply?



☐ User B's reply states a SIMILAR viewpoint to mine and I DON'T feel offended

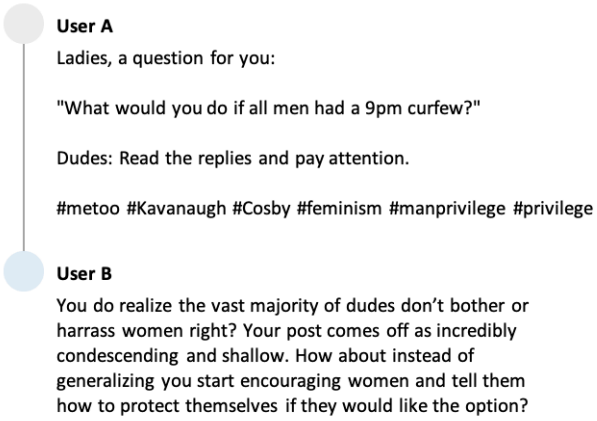
☐ User B's reply states a SIMILAR viewpoint to mine and is pretty OFFENSIVE to me

☐ User B's reply DIFFERS from my opinion and I DON'T feel offended

☐ User B's reply DIFFERS from my opinion and is pretty OFFENSIVE to me

☐ Other

3.2 Did you think User B tried to be polite?



☐ Yes

☐ No

☐ Not sure

3.3 Why do you think User B is polite? Or why do you think User B is not polite?

Answer: _____

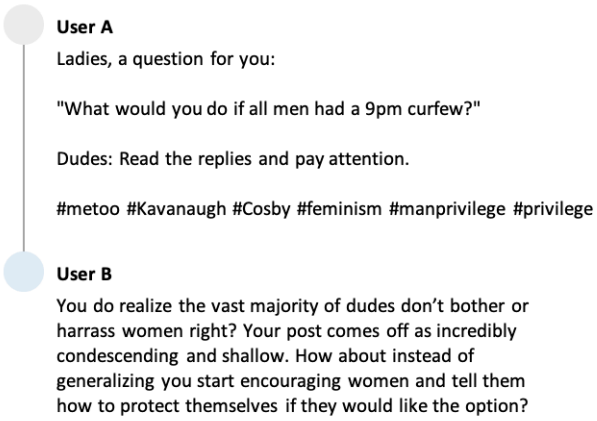
3.4 Did you try to be polite to User B?

☐ Yes

☐ No

☐ Not sure

3.5 If the conversation took place person to person, would you respond to User B as you did above?



☐ Yes

☐ No

☐ Not sure

3.6 If your answer is "NO", please provide your new answer. Otherwise, please skip this question.

Answer: _____

***Section 4**: The last task includes five questions. Please recall your experience of using the following sentence "Could you please fuck yourself" and choose the option that best describes your behaviour.*

4.1 Would you say the sentence 'Could you please fuck yourself' in face-to-face?

☐ Yes

☐ No

☐ Can't remember

4.2 Would you say the sentence 'Could you please fuck yourself' on Twitter?

☐ Yes

☐ No

☐ Can't remember

4.3 Why do you use "Could you please" in "Could you please fuck yourself"? Or why do you think others use it in that way?

Answer: _____

4.4 Do you remember using or seeing polite behaviours on Twitter which were not meant to be polite?

☐ Yes

☐ No

☐ Can't remember

4.5 If your answer is "YES", can you give me an example? Otherwise, please skip this question.

Answer: _____

Reference Lists

Adamczak-Krysztofowicz, S., & Szczepaniak-Kozak, A. (2017). A disturbing view of intercultural communication: Findings of a study into hate speech in Polish. *Linguistica Silesiana*, 38, 285–310.

Agler, D. W. (2012). *Critical Thinking - Handout 3 - The Ad Hominem Fallacy*. [Class handout]. The Pennsylvania State University. http://www.davidagler.com/teaching/criticalthinking/handouts/Handout3_AdHominemFallacy.pdf

Aijmer, K. (2015). "Will you fuck off please". The use of please by London teenagers. *Sociocultural Pragmatics*, 3(2), 127–149.

Aktaş, B., & Kohnert, A. (2020). TwiConv: A Coreference-annotated Corpus of Twitter Conversations. In M. Ogrodniczuk, V. Ng, Y. Grishina, & S. Pradhan (Eds.), *Proceedings of the Third Workshop on Computational Models of Reference, Anaphora and Coreference* (pp. 47–54). Association for Computational Linguistics. https://aclanthology.org/2020.crac-1.6

Alaghbary, G. S. (2019). Ideological positioning in conflict: the United States and Egypt' domestic political trajectory. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 83–102). Routledge.

Alharbi, A., & Lee, M. (2021). Kawarith: an Arabic Twitter Corpus for Crisis Events. In N. Habash, H. Bouamor, H. Hajj, W. Magdy, W. Zaghouani, F. Bougares, N. Tomeh, & I. A. Farha (Eds.), *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, (pp. 42–52). Association for Computational Linguistics. https://aclanthology.org/2021.wanlp-1.5

Al-Hindawi, F. H., Lect, A., & AbuKrooz, H. H. (2018). A Pragmatic Study of Argument in atimatulzahra's (PBUH) Speech. *Adab Al-Kufa*, 2(37), 45–72.

Almutairi, S. (2021). Disagreement strategies and (Im) politeness in Saudis' Twitter Communication. *Journal of Languages, Texts, and Society*, 5, 1–41.

Alorainy, W., Burnap, P., Liu, H., Javed, A., & Williams, M. L. (2018). Suspended accounts: A source of tweets with disgust and anger emotions for augmenting hate speech data sample. *Proceedings of 2018 International Conference on Machine Learning and Cybernetics* (pp. 581–586). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/ICMLC.2018.8527001

Altexsof—What is API: Definition, types, specifications, documentation. (2021). https://www.altexsoft.com/blog/engineering/what-is-api-definition-types-specifications-documentation/

Altoaimy, L. (2018). Driving change on Twitter: A corpus-assisted discourse analysis of the Twitter debates on the Saudi ban on women driving. *Social Sciences*, 7(5), 81.

Anderson, J. F., Beard, F. K., & Walther, J. B. (2010). Turn-taking and the local management of conversation in a highly simultaneous computer-mediated communication system. *Language@Internet*, 7, Article 7. http://nbn-resolving.de/urn:nbn:de:0009-7-28048

Androutsopoulos, J. (2006). Introduction: Sociolinguistics and computer-mediated communication. *Journal of sociolinguistics*, 10(4), 419–438.

Anger, I., & Kittl, C. (2011). Measuring influence on Twitter. In S. Lindstaedt, & M. Granitzer (Eds.). *i-KNOW '11: Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies* (pp. 1–4). Association for Computing Machinery. https://doi.org/10.1145/2024288.2024326

Angouri, J., & Locher, M. A. (2012). Theorising disagreement. *Journal of Pragmatics*, 44, 1549–1553.

Araque, O., Corcuera-Platas, I., Sanchez-Rada, J. F., & Iglesias, C. A. (2017). Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Systems with Application*, 77, 236–246.

Archer, D., Lansley, C., & Garner, A. (2019). The value of linguistics in assessing potential threats in an airport setting. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 454–471). Routledge.

Arcidiacono, F., & Pontecorvo, C. (2009). Cultural practices in Italian family conversations: Verbal conflict between parents and preadolescents. *European journal of psychology of education*, 24(1), 97.

Assimakopoulos, S., Baider, F. H., & Millar, S. (Eds.). (2017). *Online hate speech in the European Union: A discourse-analytic perspective*. Springer Nature.

Austin, J. L. (1962). *How to Do Things with Words*. Oxford University Press.

Aycock, A. (1995). "Technologies of the self:" Foucault and Internet discourse. *Journal of Computer-Mediated Communication*, 1(2). https://doi.org/10.1111/j.1083-6101.1995.tb00328.x

Bardovi-Harlig, K., & Salsbury, T. (2004). The organization of turns in the disagreements of L2 learners: a longitudinal perspective. In D. Boxer, & A. Cohen (Eds.), *Studying speaking to inform second language learning* (pp. 199–227). Multilingual Matters.

Baron, N. S. (1984). Computer mediated communication as a force in language change. *Visible language*, 18(2), 118–141.

Bartos, O. J., & Wehr, P. (2002). *Using conflict theory*. Cambridge University Press.

Baym, N. K. (1995). The performance of humor in computer-mediated communication. *Journal of computer-mediated communication*, 1(2). https://doi.org/10.1111/j.1083-6101.1995.tb00327.x.

Beach, A. (2019) , "It's So Bomb": Exploring Corpus-Based Threat Detection on Twitter with Discourse. *UVM Honors College Senior Theses*, 298.

Behnam, B., & Niroomand, M. (2011). An Investigation of Iranian EFL Learners' Use of Politeness Strategies and Power Relations in Disagreement across Different Proficiency Levels. *English language teaching*, 4(4), 204–220.

Bell, A., Sharma, D., & Britain, D. (2016). Labov in sociolinguistics: an introduction. *Journal of Sociolinguistics*, 20(4), 399–408.

Ben-Menachem, E. T., & Livnat, Z. (2018). Desirable and undesirable disagreements: Jewish women studying the talmudic texts. *Journal of Pragmatics*, 138, 30–44.

Benwell, B., & Stokoe, E. (2006). *Discourse and identity*. Edinburgh University Press.

Berglund, T. Ö. (2009). Disrupted turn adjacency and coherence maintenance in instant messaging conversations. *Language@Internet*, 6, Article 2. http://nbn-resolving.de/urn:nbn:de:0009-7-21066

Blitvich, P. G. C. (2018). Globalization, transnational identities, and conflict talk: the complexity of the Latino identity. *Journal of Pragmatics*, 134, 120–133.

Blitvich, P. G. C. (2019). "You are shamed for speaking it or for not speaking it good enough": the paradoxical status of Spanish in the US Latino community. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 398–416). Routledge.

Blitvich, P. G. C., & Bou-Franch, P. (2019). Introduction to analyzing digital discourse: new insights and future directions. In P. G. C. Blitvich, & P. Bou-Franch (Eds.), *Analyzing Digital Discourse* (pp. 3–22). Palgrave Macmillan.

Boberg, C. (2017). Surveys: The use of written questionnaires in sociolinguistics. In C. Mallinson, B. Childs, & G. V. Herk (Eds.), *Data Collection in Sociolinguistics: Methods*

and Applications (2nd Edition) (pp. 131–141). Routledge. (Original work published in 2013).

Bogdanov, P., Busch, M., Moehlis, J., Singh, A. K., & Szymanski, B. K. (2013). The social media genome: Modeling individual topic-specific behavior in social media. *ASONAM '13: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and* Mining (pp. 236–242). Association for Computing Machinery. https://doi.org/10.1145/2492517.2492621

Boring, N. (2020). France: Constitutional Court Strikes Down Key Provision of Bill on Hate Speech. Library of Congress. https://www.loc.gov/law/foreign-news/article/france-constitutional-court-strikes-down-key-provisions-of-bill-on-hate-speech/

Bouko, C., & Garcia, D. (2019). Citizens' reactions to Brexit on Twitter: A content and discourse analysis. In V. Koller, S. Kopf, & M. Miglbauer (Eds.), *Discourses of Brexit* (pp. 171–190). Routledge

Bousfield, D. (2007). Beginnings, middles and ends: a biopsy of the dynamic of impolite exchanges. *Journal of Pragmatics*, 39 (12), 2185–2216.

Bousfield, D. (2008). *Impoliteness in interaction*. John Benjamins Publishing Company.

Bou-Franch, P., & Blitvich, P. G. C. (2014). Conflict management in massive polylogues: A case study from YouTube. *Journal of Pragmatics*. 73, 19–36.

Bova, A., & Arcidiacono, F. (2018). The interplay between parental argumentative strategies, children's reactions and topics of disagreement during family conversations. *Learning, Culture and Social Interaction*, 19, 124–133.

Boyd, d., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. *2010 43rd Hawaii International Conference on System Sciences* (pp. 1–10). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/HICSS.2010.412

Black, L. W., & Wiederhold, A. (2014). Discursive strategies of civil disagreement in public dialogue groups. *Journal of Applied Communication Research*, 42(3), 285–306.

Briggs, C. L . (1996). Introduction. In C. L. Briggs (Eds.), *Disorderly discourse: narrative, conflict, and inequality* (pp. 3–40). Oxford University Press.

Brown, P. (1990). Gender, politeness, and confrontation in Tenejapa. *Discourse Processes*, 13, 123–141.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (2nd Edition). The Press Syndicate of the University of Cambridge. (Original work published in 1978).

Bruns, A., & Moe, H. (2014). Structural layers of communication on Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Eds.), *Twitter and society* (pp. 15–28). Peter Lang.

Bruns, A., & Burgess, J. E. (2011). The use of Twitter hashtags in the formation of ad hoc publics. In Bruns, A & De Wilde, P (Eds.), *Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011* (pp. 1–9). The European Consortium for Political Research.

Burnap, P., & Williams, M. L. (2015). Cyber hate speech on twitter: An application of machine classification and statistical modelling for policy and decision making. *Policy & Internet*, 7(2), 223–242.

CEBC—Domestic Abuse Intervention Project—The Duluth Model (DAIP). (2022). Retrieved 22nd February 2022, from https://www.cebc4cw.org/program/domestic-abuse-intervention-project/detailed

Chen, Y., Zhou, Y., Zhu, S., & Xu, H. (2012). Detecting offensive language in social media to protect adolescent online safety. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (pp. 71–80). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/SocialCom-PASSAT.2012.55

Chi, F. M. (2016). How University Students Managed Conflictual Talk in Small-Group Text Discussion. *International Journal of Applied Linguistics and English Literature*, 5(4), 154–162.

Chua, V., Madej, J., & Wellman, B. (2010). Personal communities. In P. Carrington & J. Scott (Eds.), *Handbook of social network analysis* (pp. 101–115). Sage.

Clayman, S. E. (2013). Turn-constructional units and the transition-relevance place. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 150–166). Wiley-Blackwell.

Codó, E. (2008). Interviews and Questionnaires. In W. Li & M. G. Moyer (Eds.), *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism* (pp. 158–176). Blackwell Publishing.

Coinnigh, M. M., Erivine, C., & Deeds, P. (2019). The Irish language in Belfast: the role of a language in post-conflict resolution. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 556–574). Routledge.

Çöltekin, Ç. (2019). A Corpus of Turkish Offensive Language on Social Media. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 6174–6184). Language Resources and Evaluation Conference. https://aclanthology.org/2020.lrec-1.758

Condon, S. L., & Čech, C. G. (2001). Profiling turns in interaction: Discourse structure and function. *Proceedings of the 34th Annual Hawaii International Conference on System Sciences.* Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/HICSS.2001.926501

Cooper, R. L. (1979). Sociolinguistic surveys: the state of the art. *Applied Linguistics*, 1(2), 113–128.

Coppock, L. (2005). *Politeness strategies in conversation closings*. [Unpublished manuscript]. Stanford University. Retrieved 22nd February 2022, from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.121.883&rep=rep1&type=pdf

Council of Europe. (1997). *Recommendation No. R (97) 20 of the committee of minister to member states on "hate speech".* https://www.coe.int/en/web/freedom-expression/committee-of-ministers-adopted-texts/-/asset_publisher/aDXmrol0vvsU/content/recommendation-no-r-97-20-of-the-committee-of-ministers-to-member-states-on-hate-speech-?_101_INSTANCE_aDXmrol0vvsU_viewMode=view

Cui, A., Zhang, M., Liu, Y., Ma, S., & Zhang, K. (2012). Discover breaking events with popular hashtags in twitter. *CIKM '12: Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 1794–1798). Association for Computing Machinery. https://doi.org/10.1145/2396761.2398519

Culpeper, J. (1996). Towards an Anatomy of Impoliteness. Journal of Pragmatics 25, 349-367.

Culpeper, J. (2008). Reflections on impoliteness, relational work and power. In D. Bousfield, M. A. Locher & D. Gruyter (Eds.), Impoliteness in Language: Studies on Its Interplay with Power in Theory and Practice (pp. 17-44). Mouton de Gruyter.

Culpeper, J. (2011a). *Impoliteness: Using language to cause offence*. Cambridge University Press.

Culpeper, J. (2011b). Politeness and Impoliteness. In K. Aijmer & G. Andersen (Eds.), *Pragmatics of Society* (pp. 393–438). Mouton de Gruyter.

Culpeper, J. (2015). Impoliteness Strategies. In A. Capone & J. Mey (Eds.), *Interdisciplinary Studies in Pragmatics, Culture and Society* (pp. 421–445). Springer.

Culpeper, J. (2021). Impoliteness and hate speech: compare and contrast. *Journal of Pragmatics*, 179, 4-11.

Culpeper, J., & Terkourafi, M. (2017). Pragmatic Approaches (Im)politeness. In J. Culpeper, M. Haugh, & D. Z. Kádár (Eds.), *The Palgrave Handbook of Linguistic (Im)politeness* (pp. 11-40). Springer Nature.

Culpeper, J., & Hardaker, C. (2017). Impoliteness. In J. Culpeper, M. Haugh, & D. Z. Kádár (Eds.), *The Palgrave Handbook of Linguistic (Im)politeness* (pp. 199-226). Springer Nature.

Culpeper, J., Haugh, M., & Kádár, D. Z. (2017). Introduction. In J. Culpeper, M. Haugh, & D. Z. Kádár (Eds.), *The Palgrave Handbook of Linguistic (Im)politeness* (pp. 89-118). Springer Nature.

Dadvar, M., Trieschnigg, R. B., & de Jong, F. M. (2013). Expert knowledge for automatic detection of bullies in social networks. *25th Benelux Conference on Artificial Intelligence, BNAIC 2013* (pp. 57–64). Delft University of Technology

Daft, R. L., & Lengel, R. H. (1983). Information richness: A new approach to managerial behavior and organization design. *Research in Organizational Behavior*, 6, 191–233.

Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. *Eleventh international AAAI conference on web and social media* (pp. 512–515). Association for the Advancement of Artificial Intelligence Press. https://ojs.aaai.org/index.php/ICWSM/article/view/14955

Davies, M. (2019). Stark choices and brutal simplicity: the blunt instrument of constructed oppositions in news editorials. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 44–63). Routledge.

De Bres, J., & Franziskus, A. (2019). Language ideologies in conflict at the workplace. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 433–448). Routledge.

December, J. (1997). *Notes on defining of computer-mediated communication*. Computer-Mediated Communication Magazine. https://www.december.com/cmc/mag/1997/jan/december.html

Decock, S., & Spiessens, A. (2017a). Customer complaints and disagreements in a multilingual business environment. A discursive-pragmatic analysis. *Intercultural pragmatics*, 14(1), 77–115.

Decock, S., & Spiessens, A. (2017b). Face-threatening e-mail negotiation in a multilingual business environment: a discursive analysis of refusal and disagreement strategies. In I. Kecskés, & S. Assimakopoulos (Eds.), *Current issues in intercultural pragmatics* (pp. 129–156). John Benjamins Publishing Company.

Dehcheshmeh, M. M. (2007). Specialized monolingual corpora in translation. *Translation Journal*, 11(2). http://translationjournal.net/journal/40corpus.htm

Dictionary—Boomer. (2022). Retrieved 22nd February 2022, from https://www.dictionary.com/browse/boomer

Dictionary—MEGA. (2022). Retrieved 22nd February 2022, from https://www.dictionary.com/browse/maga

Dinakar, K., Jones, B., Havasi, C., Lieberman, H., & Picard, R. (2012). Common sense reasoning for detection, prevention, and mitigation of cyberbullying. *ACM Transactions on Interactive Intelligent Systems (TiiS),* 2(3), 1–30.

Djuric, N., Zhou, J., Morris, R., Grbovic, M., Radosavljevic, V., & Bhamidipati, N. (2015). Hate speech detection with comment embeddings. *WWW'15 Companion: Proceedings of the 24th International Conference on World Wide* Web (pp. 29–30). https://doi.org/10.1145/2740908.2742760

Dobs, A., & Blitvich, P. G. C. (2013). Impoliteness in interaction: accounting for face-threat witness's responses. *Journal of Pragmatics*, 53, 112–130.

Donelan, H. (2010). Asynchronous and synchronous communication technologies. In H. Donelan, K. Kear, & M. Ramage (Eds.), *Online Communication and Collaboration* (pp. 83–88). Routledge.

Dubrovsky, V., Kiesler, S., & Sethna, B. N. (1990). Expected and unexpected effects of computer media on group decision making. *ACM SIGCHI Bulletin*, 21(3), 18–20.

Dynel, M. (2013). Being cooperatively impolite: Grice's model in the context of (im)politeness theories. In I. Kecskes, & J. Romero-Trillo (Eds.), *Research trends in intercultural pragmatics* (pp. 55–83). Mouton de Gruyter.

Eelen, G. (2001). *Critique of Politeness Theories*. St Jerome's Press.

Edstrom, A. (2004). Expressions of disagreement by Venezuelans in conversation: reconsidering the influence of culture. *Journal of Pragmatics*, 36(8), 1499–1518.

Eisenberg, A. R., & Garvey, C. (1981). Children's use of verbal strategies in resolving conflicts. *Discourse Processes*, 4, 149–170.

ElSherief, M., Kulkarni, V., Nguyen, D., Wang, W. Y., & Belding, E. (2018). Hate lingo: A target-based linguistic analysis of hate speech in social media. *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 42–51). https://ojs.aaai.org/index.php/ICWSM/article/view/15041

Emojipedia—Slightly Smiling Face. (2021). https://emojipedia.org/slightly-smiling-face/

Emoji Research Team. (2015). *2015 Emoji Report*. Emogi. http://cdn.emogi.com/docs/reports/2015_emoji_report.pdf

Eslami, Z. (2005). Invitations in Persian and English: Ostensible or Genuine?. *Intercultural Pragmatics*, 2(4), 453–480.

European Commission against Racism and Intolerance. (2015). *ECRI General Policy Recommendation N°15 on Combating Hate Speech (adopted on 8 December 2015)*. https://rm.coe.int/ecri-general-policy-recommendation-no-15-on-combating-hate-speech/16808b5b01

European Cooperation in Science and Technology. (2021). *The Dawn of the Human-Machine Era: A forecast of new and emerging language technologies*. https://ruor.uottawa.ca/bitstream/10393/42795/1/The%20LITHME%20Forecast%20report%202021.pdf

European Union. (2016). *Code of conduct on countering illegal hate speech online*. https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combatting-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en

European Union, European Commission (2019). *5th evaluation of the Code of Conduct.* https://ec.europa.eu/info/sites/info/files/codeofconduct_2020_factsheet_12.pdf

Fernández-Amaya, L. (2019). Disagreement and (im)politeness in a Spanish family members' WhatsApp group. *Russian Journal of Linguistics*, 23, 1065–1087.

Fraser, B. (1990). Perspectives on politeness. *Journal of Pragma*tics, 14, 219–236.

Gal, S., & Irvine, J. T. (2019). *Signs of Difference: Language and Ideology in Social Life*. Cambridge University Press.

Garcia, A. C., & Jacobs, J. B. (1998). The interactional organization of computer mediated communication in the college classroom. *Qualitative Sociology,* 21(3), 299–317.

Garcia, A. C., & Jacobs, J. B. (1999). The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on language and social interaction*, 32(4), 337–367.

Garton, L. E., & Wellman, B. (1995). Social impacts of electronic mail in organizations: A review of the research literature. *Annals of the International Communication Association*, 18(1), 434–453.

Georgakopoulou, A. (2001). Arguing about the future: On indirect disagreements in conversations. *Journal of Pragmatics*, 33(12), 1881–1900.

Giarolo, K. A. (2018). Grice on rationality. *Filosofia Unisinos*, 19(3), 223–234.

Gibbs, R. W. (1999). *Intentions in the Experience of Meaning*. Cambridge University Press.

Gibson, W. (2009). Intercultural communication online: Conversation analysis and the investigation of asynchronous written discourse. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 10(1). https://doi.org/10.17169/fqs-10.1.1253

Giles, D., Stommel, W., Paulus, T., Lester, J., & Reed, D., (2015). Microanalysis of online data: The methodological development of "digital CA". *Discourse, Context and Media*, 7, 45-51.

Gitari, N. D., Zuping, Z., Damien, H., & Long, J. (2015). A lexicon-based approach for hate speech detection. *International Journal of Multimedia and Ubiquitous Engineering*, 10(4), 215–230.

Gleason, B. (2016). New literacies practices of teenage Twitter users. Learning, *Media and Technology*, 41(1), 31–54.

Goffman, E. (1955). On Face-Work: An Analysis of Ritual Elements in SocialInteraction. *Psychiatry*, 18 (3), 213–231.

Goffman. E. (1967). *Interaction Ritual; Essays on Face-to-Face Behaviour*. Doubleday.

Goodwin, M. H., & Goodwin. C. (1987). Children's arguing. In S. U. Philips, S. Steele, & C. Tanz (Eds.), *Language, Gender and Sex. In Comparative Perspective* (pp. 200–248). Cambridge University Press.

Goodwin, C. & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology*, 19, 283–307.

Graham, S. (2007). Disagreeing to agree: Conflict, (im)politeness and identity in a computer-mediated community. *Journal of pragmatics*, 39(4), 724-759.

Graham, S. L. (2019). Interaction and conflict in digital communication. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 310–328). Routledge.

Grannan, C. (n.d.). *What's the Difference Between the Deep Web and the Dark Web?*. Britannica. https://www.britannica.com/story/whats-the-difference-between-the-deep-web-and-the-dark-web

Grice, P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp. 41–58). Academic Press.

Grice, P. (1989). *Studies in the way of words*. Harvard University Press.

Grimminger, L., & Klinger, R. (2021). *Hate Towards the Political Opponent: A Twitter Corpus Study of the 2020 US Elections on the Basis of Offensive Speech and Stance Detection.* arXiv. https://doi.org/10.48550/arXiv.2103.01664

Grimshaw, A. (1990). Introduction. In A. Grimshaw (Ed.), *Conflict Talk: Sociolinguistic Investigations of Arguments in Conversations* (pp. 1–20). Cambridge University Press.

Gruber, H. (1998). Computer-mediated communication and scholarly discourse: Forms of topic-initiation and thematic development*. Pragmatics*, 8(1), 21–45.

Gruber, H. (2017). Quoting and retweeting as communicative practices in computer mediated discourse. *Discourse, context & media*, 20, 1–9.

Gruzd, A., Wellman, B., & Takhteyev, Y. (2011). Imagining Twitter as an imagined community. *American Behavioral Scientist*, 55(10), 1294–1318.

Gu, Y. (1990). Politeness Phenomena in Modern Chinese. *Journal of Pragmatics*, 14(2), 237–258.

Habib, R. (2008). Humor and disagreement: Identity construction and cross-cultural enrichment. *Journal of Pragmatics,* 40, 1117–1145.

Halavais, A. (2014). Structure of Twitter: Social and Technical. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Eds.), *Twitter and society* (pp. 29–42). Peter Lang.

Hamdi, S. A. (2021). Mining ideological discourse on Twitter: The case of extremism in Arabic. *Discourse & Communication*. https://doi.org/10.1177/17504813211043706

Haralambos, M., & Holborn, M. (1991). *Sociology: Themes and Perspectives* (3rd Edition). Collins Educational. (Original work published in 1980).

Holmes, J., & and Schnurr, S. (2005). Politeness, Humour and Gender in the Workplace: Negotiating Norms and Identifying Contestation. *Journal of Politeness Research*, 1 (1), 121–149.

Harris, Z. S. (1952). Discourse Analysis: A Sample Text. *Language*, 28(4), 474-494.

Harrison, H. L., Kochalski, S., Arlinghaus, R., & Aas, Ø. (2019). 'Do you care about the river?' A critical discourse analysis and lessons for management of social conflict over Atlantic salmon (Salmo salar) conservation in the case of voluntary stocking in Wales. *People and Nature*, 1(4), 507–523.

Harvey, R. (2020). Twitter reactions to the UN's #HeForShe campaign for gender equality: A corpus-based discourse analysis. *Journal of Corpora and Discourse Studies*, 3, 31–50.

Haugh, M. (2008). Intention in pragmatics. *Intercultural Pragmatics*, 5 (2), 99-110.

Haugh, M. (2022). (Online) public denunciation, public incivilities and offence. *Language & Communication*, 87, 44-59.

Haugh, M., & Culpeper, J. (2018). Integrative pragmatics and (im)politeness theory. In C. Ilie & N. R. Norrick (Ed.), Pragmatics and its interfaces (pp. 213-239). John Benjamins Publishing Company.

Haugh, M., & Sinkeviciute, V. (2019). Offence and conflict talk. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 196–214). Routledge.

Hayashi, M. (2013). Turn Allocation and Turn Sharing. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 167–190). Wiley-Blackwell.

Herring, S. (1994). Politeness in computer culture: Why women thank and men flame. In M. Bucholtz, A. C. Liang, Laurel A. Sutton, Caitlin Hines (Eds.), *Culture performances. Proceedings of the Third Berkeley Women and Language Conference* (pp. 278–294), Berkeley.

Herring, S. C. (1996a). Introduction. In S. C. Herring, (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 1–12). John Benjamins Publishing.

Herring, S. (1996b). Posting in a different voice: Gender and ethics in computer-mediated communication. In C. Ess (Ed.), *Philosophical Perspectives in Computer-Mediated Communication* (pp. 115–145). State University of New York Press.

Holmes, J., & and Schnurr, S. (2005). Politeness, Humour and Gender in the Workplace: Negotiating Norms and Identifying Contestation. *Journal of Politeness Research*, 1 (1), 121–149.

Honda, A. (2002). Conflict management in Japanese public affairs talk shows. Journal of *Pragmatics*, 34(5), 573–608.

Honeycutt, C., & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. *2009 42nd Hawaii International Conference on System Sciences* (pp. 1–10). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/HICSS.2009.89

Housley, W., Webb, H., Edwards, A., Procter, R., & Jirotka, M. D. A. (2017). Digitizing Sacks? Approaching social media as data. *Qualitative Research*, 17(6), 627–644.

Hovy, E., & Lavid, J. (2010). Towards a 'science' of corpus annotation: a new methodological challenge for corpus linguistics. *International journal of translation*, 22(1), 13–36.

Hutchby, I. (1992). Confrontation talk: aspects of 'interruption' in argument sequences on talk radio. *Text*, 12(3), 343–371.

Hutchby, I. (2001a). Confrontation as a Spectacle: The Argumentative Frame of the Jerry Springer Show. In A. Tolson (Ed.), *Television Talk Shows: Discourse, Performance, Spectacle* (pp. 155–172). Lawrence Erlbaum Associate.

Hutchby, I. (2001b). *Conversation and technology.* Polity Press.

Ilbury, C. (2020). "Sassy Queens": Stylistic orthographic variation on Twitter and the enregisterment of AAVE. *Journal of Sociolinguistics*, 24(2), 245–264.

Ide, S. (1989). Formal Forms and Discernment: Two Neglected Aspects of Universals of Linguistic Politeness. *Multilingua*, 8, 223–248.

Iman, Z., Sanner, S., Bouadjenek, M. R., & Xie, L. (2017). A Longitudinal Study of Topic Classification on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, 11(1), 552–555.

Imran, M., Mitra, P., & Castillo, C. (2016). *Twitter as a lifeline: Human-annotated Twitter corpora for NLP of crisis-related messages*. arXiv. https://arxiv.org/abs/1605.05894

Ishihara, N., & Prado, M. C. D. A. (2021). The Negotiation of Meaning in Aviation English as a Lingua Franca: A Corpus-Informed Discursive Approach. *The Modern Language Journal*, 105(3), 639–654.

Iveson, M. (2017). Gendered dimensions of Catalan nationalism and identity construction on Twitter. *Discourse & Communication*, 11(1), 51–68.

James, W. (2019). *Instagram, Facebook and Twitter struggle to contain the epidemic in online bullying*. MarktWatch. https://www.marketwatch.com/story/why-it-may-be-too-late-for-instagram-facebook-and-twitter-to-contain-the-epidemic-in-online-bullying-2019-07-15

Jary, M. (1998). Relevance theory and the communication of politeness. *Journal of Pragmatics*, 30, 1-19.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter: understanding microblogging usage and communities. *WebKDD/SNA-KDD '07: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis* (pp. 56-65). Association for Computing Machinery. https://doi.org/10.1145/1348549.1348556

Jeffries, L., & O'Driscoll, J. (2019). Introduction: the origins of The Routledge Handbook of Language in Conflict. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 1–11). Routledge.

Jenks, C., & Firth, A. (2013). Multi-party voice-based chat room interaction. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *Handbook of pragmatics of computer-mediated communication* (pp. 217–241). De Gruyter.

Jeong, H. W. (2008). *Understanding conflict and conflict analysis*. Sage.

Jeon, M., Jun, S., & Hwang, E. (2014). Hashtag recommendation based on user tweet and hashtag classification on twitter. In Y. G. Chen, W. T. Balke, J. L. Xu, W. Xu, P. Q. Jin, X. Lin, T. Tang, & E. Hwang (Eds.), *Web-Age Information Management* (pp. 325–336). Springer. https://doi.org/10.1007/978-3-319-11538-2_30

Johnstone, B. (1989). Linguistic strategies and cultural styles for persuasive discourse. In S. Ting-Toomey, & F. Korzenny (Eds.), *Language, Communication, and Culture: Current Directions* (pp. 139–156). Sage.

Johnstone, B. (2018). *Discourse Analysis* (3rd Edition). Oxford: Blackwell. (Original work published in 2002).

Johnson, R. T., & Johnson, D. W. (1985). Student-student interaction: Ignored but powerful. *Journal of teacher education*, 36(4), 22–26.

Jonathan, V. (2017). *Here's how Twitter is trying to stop bulling and abuse on its service*. Fortune. https://fortune.com/2017/02/07/twitter-bullying-abuse-service/

Jones, R. H. (2019). Discourse Analysis: A resource book for students (2nd Edition). Routledge. (Original work published in 2012).

Jones, T. (2015). Toward a description of African American Vernacular English dialect regions using "Black Twitter". *American Speech,* 90, 403–440.

Kakavá, C. (2002). Opposition in modern Greek discourse: cultural and contextual constraints. *Journal of Pragmatics*, 34, 1537–1368.

Kansara, K. B., & Shekokar, N. M. (2015). A framework for cyberbullying detection in social network. *International Journal of Current Engineering and Technology*, *5*(1), 494–498.

Kennedy, G. (2014). *An introduction to corpus linguistics*. Routledge.

Khammari, H. (2021). The realization of the speech act of disagreement by Tunisian non-native and American native speakers of English. *Arab Journal of Applied Linguistics*, 6(1).

Khan, K. (2019). Hate crimes: language, vulnerability and conflict. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 417–432). Routledge.

Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, 39 (10), 1123–1134.

Kiesler, S., Zubrow, D., Moses, A. M., & Geller, V. (1985). Affect in computer-meditated communication: An experiment in synchronous terminal-to-terminal discussion. *Human-Computer Interaction*, 1(1), 77–104.

Kim, S., Weber, I., Wei, L., & Oh, A. (2014). Sociolinguistic analysis of Twitter in multilingual societies. *HT '14: Proceedings of the 25th ACM conference on Hypertext and social media* (pp. 243–248). Association for Computing Machinery. https://doi.org/10.1145/2631775.2631824

Klein, L. E. (2009). Liberty, manners and politeness in early eighteenth-century England. *The Historical Journal*, 32 (3), 583–605.

Koczogh, H. V. (2012). The role of gender in verbal disagreement: A study of disagreement strategies employed by Hungarian undergraduate students. *Gender Studies*, 11, 318–332.

Koester, A. (2017). Conflict talk. In B. Vine (Ed.), *The Routledge handbook of language in the workplace* (pp. 272–283). Routledge.

Koutlaki, S. (2002). Offers and Expressions of Thanks as Face-Enhancing Acts: tae'rof in Persian. *Journal of Pragmatics*, 34(12), 1733–1756.

Kwok, I., & Wang, Y. (2013). Locate the hate: Detecting tweets against blacks. *Proceedings of the Twenty-Seventy AAAI Conference on Artificial Intelligence* (pp. 1621–1622). Association for the Advancement of Artificial Intelligence. https://www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/view/6419/6821

Lachenicht, L. G. (1980). Aggravating Language: A study of abusive and insulting language. *International Journal of Human Communication*, 13(4), 607-688.

Lakoff, R. T. (1973). The Logic of Politeness; or Minding You P's and Q's. In C. Corum, T. C. Smith-Start, & A. Weiser (Eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society* (pp. 292-305). Chicago Linguistic Society.

Lebra, T. S. (1976). *Japanese patterns of behavior*. Honolulu: University of Hawaii Press.

Lee, G. B. (1996). Addressing anonymous messages in cyberspace. *Journal of Computer-Mediated Communication*, 2(1). https://doi.org/10.1111/j.1083-6101.1996.tb00183.x

Lee, J. S. (2020). Sports, Instagram, and conflict talk in Englishes. *World Englishes*, 39, 94–105.

Lee, K., Palsetia, D., Narayanan, R., Patwary, M. M. A., Agrawal, A., & Choudhary, A. (2011). Twitter trending topic classification. *2011 IEEE 11th International Conference on Data Mining Workshops* (pp. 251–258). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/ICDMW.2011.171

Leech, G. N. (1977). *Language and Tact*, Series A (46). Linguistic Agency, University of Trier.

Leech, G. N. (2014). *The Pragmatic of Politeness*. Oxford University Press. (Original work published in 1983).

Leech, G. & Thomas, J. (1990). Language, meaning and context. In N. E. Collinge (Ed.), *An encyclopedia of language* (pp. 173–206). Routledge.

Levisen, C. (2016). The ethnopragmatics of speech acts in postcolonial discourse: “truth” and “trickery” in a South Pacific tale. In C. Schubert, & L. Volkmann (Eds.), *Pragmatic Perspectives on Postcolonial Discourse: Linguistics and Literature* (pp. 41–64). Cambridge Scholars Publishing.

Levisen, C. (2019). Biases we live by: Anglocentrism in linguistics and cognitive sciences. *Language Sciences,* 76, 101–173.

Licoppe, C., & Morel, J. (2012). Video-in-interaction: “Talking heads” and the multimodal organization of mobile and Skype video calls. *Research on Language and Social Interaction*, 45(4), 399–429.

Licoppe, C., & Morel, J. (2018). Visuality, text and talk, and the systematic organization of interaction in Periscope live video streams. *Discourse Studies*, 20(5), 637–665.

Li, J., Xu, H., He, X., Deng, J., & Sun, X. (2016). Tweet modeling with LSTM recurrent neural networks for hashtag recommendation. *2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 1570–1577). Institute of Electrical and Electronics Engineers. http://doi.org/ 10.1109/IJCNN.2016.7727385

Liu, G., Amini, M. H., Boroojeni, K. G., Nejadpak, A., & Iyengar, S. S. (2016). Best practices for online marketing in twitter: an experimental study. *2016 IEEE International Conference on Electro Information Technology (EIT)* (pp. 504–509). Institute of Electrical and Electronics Engineers. https://ieeexplore.ieee.org/document/7535292

Locher, M. A. (2004). *Power and politeness in action: disagreements in oral communication*. Mouton de Gruyter.

Locher, M. A. (2006a). Polite Behaviour Within Relational Work: The Discursive approach to politeness. *Multilingual*, 25(3), 249-267.

Locher, M. A. (2006b). The Discursive Approach to Polite Behaviour. *Language in Society*, 35, 733-735.

Locher, M. A. & R. J. Watts (2005). Politeness theory and relational work. *Journal of Politeness Research: Language, Behaviour, Culture*, 1(1), 9-33.

Locher, M. A. & R. J. Watts (2008). Relational work and impoliteness: Negotiating norms of linguistic behaviour. In D. Bousfield & M. A. Locher (Eds.), *Impoliteness in language: Studies on its interplay with power in theory and practice* (pp. 77-99). Mouton de Gruyter.

Loos, E. E., Anderson, S., Day, D. H., Jordan, P. C., & Wingate, J. D. (2013). Turn. In *SIL Glossary of Linguistic Terms*. https://glossary.sil.org/term/turn

Lundberg, J., & Laitinen, M. (2020). Twitter trolls: a linguistic profile of anti-democratic discourse. *Language Sciences*, 79, 1–14.

Ma, Z., Sun, A., & Cong, G. (2012). Will this# hashtag be popular tomorrow?. *SIGIR '12: Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (pp. 1173–1174). Association for Computing Machinery. https://doi.org/10.1145/2348283.2348525

Mao, Z. D. (1990). *Mao Zedong on Dialectical Materialism: Writings on Philosophy, 1937*. Routledge.

Marcinowski, M., & Ławrynowicz, A. (2020). Predicting the outbreak of conflict in online discussions using emotion-based features. In M. Bielikova, T. Mikkonen & C. Pautasso (Eds.), *Web Engineering: 20th International Conference, ICWE 2020, Helsinki, Finland, June 9–12, Proceedings* (pp. 505–511). Springer. https://doi.org/10.1007/978-3-030-50578-3_35

Masud, S., Dutta, S., Makkar, S., Jain, C., Goyal, V., Das, A., & Chakraborty, T. (2021). *Hate is the New Infodemic: A Topic-aware Modeling of Hate Speech Diffusion on Twitter*. arXiv. https://arxiv.org/abs/2010.04377v1

Matsumoto, Y. (1989). Politeness and Conversational Universals–Observations from Japanese. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, 8(2–3), 207–222.

McArthur, T., Lam-McArthur, J., & Fontaine, L. (Eds.) (2018). Computer-Mediated Communication. In *The Oxford Companion to the English Language* (2nd Edition). Oxford University Press. (Original work published in 2003). https://www.oxfordreference.com/view/10.1093/acref/9780199661282.001.0001/acref-9780199661282

McEnery, T., & Hardie, A. (2013). The history of corpus linguistics. In K. Allan (Ed.), *The Oxford handbook of the history of linguistics* (pp. 727–745). Oxford University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Taylor & Francis.

McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. McGraw Hill.

Meredith, J. (2017). Analysing technological affordances of online interactions using conversation analysis. *Journal of Pragmatics*, 115, 42–55.

Meredith, J (2019). Conversation analysis on online interaction. *Research on Language and Social Interaction*, 52(3), 241-256.

Merriam-Webster—Stream. (2022). Retrieved 22nd February 2022, from https://www.merriam-webster.com/dictionary/stream#h2

Merriam-Webster—Streamer. (2022). Retrieved 22nd February 2022, from https://www.merriam-webster.com/dictionary/streamer

Merriam-Webster—Lady. (2022). Retrieved 22nd February 2022, from https://www.merriam-webster.com/dictionary/lady

Mesthrie, R. (2009). Cleaning the Ground: Basic issues, concepts and approaches. In R. Mesthrie, J. Swann, A. Deumert, & W. L. Leap (Eds.), *Introducing* Sociolinguistics (2nd Edition) (pp. 1–41). Edinburgh University Press. (Original work published in 2000).

Meyerhoff, M. (2011). *Introducing sociolinguistics* (2nd Edition). Routledge. (Original work published in 2006).

Miller-Ott, A. E., Kelly, L., & Duran, R. L. (2012). The effects of cell phone usage rules on satisfaction in romantic relationships. *Communication Quarterly*, 60(1), 17–34.

Mills, S. (2003). *Gender and politeness.* Cambridge University Press.

Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and Interpretation*. Blackwell.

Moernaut, R., Mast, J., Temmerman, M., & Broersma, M. (2020). Hot weather, hot topic. Polarization and sceptical framing in the climate debate on Twitter. *Information, Communication & Society*, 1–20. https://doi.org/10.1080/1369118X.2020.1834600

Mondal, M., Silva, L. A., & Benevenuto, F. (2017). A measurement study of hate speech in social media. *HT'17: Proceedings of the 28th ACM Conference on Hypertext and Social Media* (pp. 85–94). Association for Computing Machinery. https://doi.org/10.1145/3078714.3078723

Muntigl, P., & Turnbull, W. (1998). Conversational structure and facework in arguing. *Journal of Pragmatics*, 29, 225–56.

Murphy, J. (2019). I'm sorry you are such an arsehole: (non-)canonical apologies and their implications for (im) politeness. *Journal of pragmatics*, 142, 223–232.

Murthy, D. (2013). *Twitter: Social Communication in the Twitter Age*. Polity Press.

Myers, D. (1987). "Anonymity is part of the magic": Individual manipulation of computer-mediated communication contexts. *Qualitative Sociology*, 10, 251–266.

Nahajec, L. (2019). Projecting your "opponent's" views: linguistic negation and the potential for conflict. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 64-82). Routledge.

Nelson, C. K. (2001). If it sounds too good to be true, it is: a Wittgensteinian approach to the conflict literature. *Language & Communication*, 21, 1–22.

Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2013). "How old do you think I am?" A study of language and age on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, 7(1), 439–448.

Nguyen, H. T. (2011). Boundary and alignment in multiparty conflict talk. *Journal of Pragmatics,* 43(6), 1755–1771.

Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). Abusive language detection in online user content. *WWW'16: Proceeding of the 25th International Conference on World Wide Web* (pp. 145–153). International World-Wide Web Conference Steering Committee. https://doi.org/10.1145/2872427.2883062

Norrick, N. R., & Spitz, A. (2008). Humor as a resource for mitigating conflict in interaction. *Journal of Pragmatics*, 40, 1661–1686.

Novak, P., Smailovic, J., Sluban, B., & Mozetic, I. (2015). Sentiment of emojis. *PLOS ONE*, 10(12).

Nurfikri, F. S., & Mubarok, M. S. (2018). News topic classification using mutual information and bayesian network. *2018 6th International Conference on Information and Communication Technology (ICoICT)* (pp. 162–166). Institute of Electrical and Electronics Engineers. 10.1109/ICoICT.2018.8528806

Nwoye, O. G. (1992). Linguistic Politeness and Sociocultural Variation of the Notion of Face. *Journal of Pragmatics*, 18(4), 309–328.

O'Driscoll, J. (2017). Face and (Im)politeness. In J. Culpeper, M. Haugh, & D. Z. Kádár (Eds.), The Palgrave Handbook of Linguistic (Im)politeness (pp. 89-118). Springer Nature.

O'Driscoll, J. (2019). Introduction: conflict as if happens. In M. Evans, L. Jeffries & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 167–175). Routledge.

Oz, M., Zheng, P., & Chen, G. M. (2018). Twitter versus Facebook: Comparing incivility, impoliteness, and deliberative attributes. *New media & society*, 20(9), 3400–3419.

Pagliai, V. (2010). Introduction: performing disputes. *Journal of Linguistic Anthropology*, 20, 63–71.

Paltridge, B. (2012). Discourse Analysis: An introduction (2nd Edition). Bloomsbury. (Original work published in 2006).

Pandharipande, R. (1992). Defining Politeness in Indian English. *World Englishes*, 11(2/3), 241–250.

Park, J. H., & Fung, P. (2017). *One-step and two-step classification for abusive language detection on Twitter*. arXiv. https://arxiv.org/abs/1706.01206

Parks, M. R., & Floyd, K. (1996). Making Friends in Cyberspace. *Journal of Computer-Mediated Communication*, 1(4). https://doi.org/10.1111/j.1083-6101.1996.tb00176.x

Patra, B. G., Debbarma, K., Das, D., & Bandyopadhyay, S. (2012). Part of Speech (POS) Tagger for Kokborok. *Proceedings of COLING 2012: Posters* (pp. 923–932). Association for Computational Linguistics. https://aclanthology.org/C12-2090.pdf

Patrick, P. L. (2004). The Speech Community. In J. K. Chambers, P. Trudgill & N. Schilling-Estes (Eds), *The Handbook of Language Variation and Change* (pp. 573–597). Blackwell Publishers.

Pearce, W. B., & Littlejohn, S. W. (1997). *Moral conflict: when social worlds collide*. Sage.

Peddinti, S. T., Ross, K. W., & Cappos, J. (2017). User anonymity on twitter. *IEEE Security & Privacy,* 15(3), 84–87.

Pennycook, A. (2017). *Posthumanist applied linguistics*. Taylor & Francis Group.

Pereira-Kohatsu, J. C., Quijano-Sánchez, L., Liberatore. F., & Camacho-Collados, M. (2019). Detecting and Monitoring Hate Speech in Twitter. Sensors, 19(21), 4654.

Pflugmacher, J., Escher, S., Reubold, J., & Strufe, T. (2020). The German-Speaking Twitter Community Reference Data Set. *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 1172–1177). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/INFOCOMWKSHPS50562.2020.9162728

Pfister, J. (2010). Is there a need for a maxim of politeness?. *Journal of Pragmatics*, 42(5), 1266–1282.

Plutis, D. (2020). *How to spot a bot on social media*. Avira. https://www.avira.com/en/blog/how-to-spot-a-bot-on-social-media

Posch, L., Wagner, C., Singer, P., & Strohmaier, M. (2013). Meaning as collective use: predicting semantic hashtag categories on twitter*. WWW '13 Companion: Proceedings of the 22nd International Conference on World Wide Web* (pp. 621–628). Association for Computing Machinery. https://doi.org/10.1145/2487788.2488008

Postmes, T., Spears, R., & Lea, M. (1999). Social identity, normative content, and "deindividuation" in computer-mediated groups. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 164–183). Blackwell Science.

Prostko, J. (2020). *Online anonymity: effects, issues, and limitations.* http://doi.org/10.13140/RG.2.2.20411.23841

Putnam, L. L., & Poole, M. S. (1987). Conflict and negotiation. In F. M. Jablin, L. L. Putnam, K. H. Roberts, & L. W. Porter (Eds.), *Handbook of organisational communication* (pp. 549–599). Sage.

Putnam, L. L. (2001). The language of opposition: challenges in organizational dispute resolution. In W. Eadie & P. Nelson (Eds.) *The language of conflict and resolution. Thousand Oaks* (pp. 9–20). Sage.

Qian, J., ElSherief, M., Belding, E. M., & Wang, W. Y. (2018). Leveraging intra-user and inter-user representation learning for automated hate speech detection. arXiv. https://arxiv.org/abs/1804.03124.

Quercia, D., Askham, H., & Crowcroft, J. (2012). Tweetlda: supervised topic classification and link prediction in twitter. *WebSci '12: Proceedings of the 4th Annual ACM Web Science Conference* (pp. 247–250). Association for Computing Machinery. https://doi.org/10.1145/2380718.2380750

Quijano-Sánchez, L., Liberatore, F., Camacho-Collados, J., & Camacho-Collados, M. (2018). Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police. *Knowledge-Based System*, 149, 155–168.

Quora—Do you like being called a bitch during sex. (2022). Retrieved 22nd February 2022, from https://www.quora.com/Do-you-like-being-called-a-bitch-during-sex/

Regulating Anti-Discrimination and Anti-Racism. (2015). *Project's Flyer*. http://win.radar.communicationproject.eu/web/wp-content/uploads/2015/01/Last-exec-of-depliant.pdf

Rajadesingan, A., Zafarani, R., & Liu, H. (2015). Sarcasm Detection on Twitter: A behavioral modelling approach. *WSDM'15: Proceedings of the 8th ACM International Conference on Web Search and Data Mining* (pp. 97–196). Association for Computing Machinery. https://doi.org/10.1145/2684822.2685316

Reed, D. J. (2017). Performance and interaction on soundcloud: Social remix and the fundamental techniques of conversation. *Journal of Pragmatics*, 115, 82–98.

Rees-Miller, J. (2000). Power, severity, and context in disagreement. *Journal of Pragmatics,* 32, 1087–1111.

Reid, E. (1991). *Electropolis: Communication and community on internet relay chat*. http://www.aluluei.com/electropolis.htm

Reyes-Menendez, A., Saura, J. R., & Thomas, S. B. (2020). Exploring key indicators of social identity in the# MeToo era: Using discourse analysis in UGC. *International Journal of Information Management*, 54. https://doi.org/10.1016/j.ijinfomgt.2020.102129

Rice, R. E. (1984). Mediated Group Communication. In R. E. Rice, & J. H. Bair (Eds.), *The new media: Communication, research, and technology*. SAGE.

Rintel, S. (2013). Video calling in long-distance relationships: The opportunistic use of audio/video distortions as a relational resource. *The Electronic Journal of Communication/La Revue Electronic De Communication (EJC/REC)*, 23(1–2). https://www.microsoft.com/en-us/research/publication/video-calling-in-long-distance-relationships-the-opportunistic-use-of-audiovideo-distortions-as-a-relational-resource/

Rintel, E. S., Mulholland, J., & Pittam, J. (2001). First things first: Internet Relay Chat openings. *Journal of Computer-Mediated Communication*, 6(3). https://doi.org/10.1111/j.1083-6101.2001.tb00125.x

Rintel, E. S., & Pittam, J. (1997). Strangers in a strange land: Interaction management on Internet Relay Chat. *Human Communication Research*, 23(4), 507–534.

Robertson, A. M., & Willett, P. (1998). Applications of n-grams in textual information systems. *Journal of Documentation*, 54(1), 48–69.

Robinson, J. D. (2013). Overall Structural Organization. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 257–280). Wiley-Blackwell.

Romero, D. M., Meeder, B., & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. *WWW '11: Proceedings of the 20th international conference on World wide web* (pp. 695–704). Association for Computing Machinery. https://doi.org/10.1145/1963405.1963503

Rosa, H., Batista, F., & Carvalho, J. P. (2014). Twitter topic fuzzy fingerprints. *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 776–783). Institute of Electrical and Electronics Engineers. http://doi.org/ 10.1109/FUZZ-IEEE.2014.6891781

Rosé, C. P. (2017). A social spin on language analysis. *Nature*, 545(7653), 166–167.

Ross, A. S., & Caldwell, D. (2020). 'Going negative': An appraisal analysis of the rhetoric of Donald Trump on Twitter. *Language & communication*, 70, 13–27.

Sacks, H. (2004). An initial characterization of the organization of speaker turn-taking in conversation. In G. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 35–42). John Benjamins.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn taking for conversation. *Language*, 50, 696–735.

Sahlane, A. (2019). Discursive (re)construction of the prelude to the 2003 Iraq War in op/ed press: dialectics of argument and rhetoric. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 25–43). Routledge.

Sahlgren. M., Isbister, T. & Olsson, F. (2018). Learning representations for detecting abusive language. *Proceedings of the 2nd workshop on abusive language online (ALW2*) (pp. 115–123). Association for Computational Linguistics. http://doi.org/ 10.18653/v1/W18-5115

Sahın, E., Aydos, M., & Orhan, F. (2018). Spam/ham e-mail classification using machine learning methods based on bag of words technique. *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1–4). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/SIU.2018.8404347

Saito, K., & Masuda, N. (2013). Two types of Twitter users with equally many followers. *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)* (pp. 1425–1426). Institute of Electrical and Electronics Engineers. http://doi.org/10.1145/2492517.2492575

Saleem, H. M., Dillon, K. P., Benesch, S., & Ruths, D. (2017). *A Web of Hate: tackling hateful speech in online social spaces.* arXiv. https://arxiv.org/abs/1709.10159

Salminen, J., Almerekhi, H., Milenković, M., Jung, S. G., An, J., Kwak, H., & Jansen, B. J. (2018a). Anatomy of online hate: developing a taxonomy and machine learning models for identifying and classifying hate in online news media. *Proceedings of the*

Twelfth International AAAI Conference on Web and Social Media (pp. 330–339). Association for the Advancement of Artificial Intelligence. https://ojs.aaai.org/index.php/ICWSM/article/view/15028

Salminen, J., Luotolahti, J., Almerekhi, H., Jansen, B. J., & Jung, S. G. (2018b). Neural network hate deletion: developing a machine learning model to eliminate hate from online comments. In S. Bodrunova (Ed.), *Internet Science: 5th International Conference, INSCI 2018, St. Petersburg, Russia, October 24-26, 2018 Proceedings* (pp. 25-39). Springer. https://doi.org/10.1007/978-3-030-01437-7_3

Salminen, J., Hopf, M., Chowdhury, S. A., Jung, S. G., Almerekhi, H., & Jansen, B. J. (2020). Developing an online hate classifier for multiple social media platforms. *Human-centric Computing and Information Sciences,* 10(1), 1–34.

Sbisà, M. (2009). Speech act theory. In J. Verschueren & J. O. Östman (Eds.), *Key Notions for Pragmatics* (pp. 229-244). John Benjamins Publishing Company.

Scheffler, T. (2014). A German Twitter Snapshot. *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14)* (pp. 2284–2289). European Language Resources Association. http://www.lrec-conf.org/proceedings/lrec2014/pdf/1146_Paper.pdf

Scheffler, T. (2017). Conversations on twitter. In D. Fišer & M. Beißwenger (Eds.), *Investigating Computer-Mediated Communication: Corpus-based approaches to language in the digital world* (pp. 124–144). Ljubljana University Press.

Schegloff, E. A. (2007). *Sequence organization in interaction: a primer in conversation analysis*. Cambridge University Press.

Schiffrin, D. (1984). Jewish argument as sociability. *Language in Society*, 13, 311–335.

Schiffrin, D. (1985). Everyday argument: The organization of diversity in talk. In T. van Dijk (Ed.), *Handbook of Discourse Analysis: Discourse and Dialogue* (pp. 35–46). Academic Press.

Schmitt, C. J., & Reiter, R. M. (2019). Leadership in conflict: Disagreement and conflict in a start-up team. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 286–307). Routledge.

Schönfeldt, J., & Golato, A. (2003). Repair in chats: A conversation analytic approach. *Research on language and social interaction*, 36(3), 241–284.

Scott, S. (2002). Linguistic feature variation within disagreements: an empirical investigation. *Text*, 22(2), 301–328.

Searle, J. (1969). *Speech Acts: An essay in philosophy of language*. Cambridge: Cambridge University Press.

Shahrokhi, M., & Bidabadi, F. S. (2013). An overview of politeness theories: Current status, future orientations. *American Journal of Linguistics*, 2(2), 17–27.

Shum, W., & Lee, C. (2013). (Im)politeness and disagreement in two Hong Kong Internet discussion forums. *Journal of Pragmatics*, 50, 52–83.

Siapera, E., Moreo, E., & Zhou, J. (2018). *Hate Track: Tracking and Monitoring Racist Speech Online*. Irish Human Rights and Equality Commission.

Siegel, J., Dubrovsky, V., Kiesler, S., & McGuire, T. W. (1986). Group processes in computer-mediated communication. *Organizational behavior and human decision processes*, 37(2), 157–187.

Sidnell, J. (2010). *Conversation Analysis: An introduction*. Wiley-Blackwell.

Sifianou, M. (2019). Conflict, disagreement and (im)politeness. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 176–195). Routledge

Simonin, O. (2018). Verbal irony, politeness… and three ironic types. In M. Jobert & S. Sorlin (Eds), *The Pragmatics of Irony and Banter* (pp. 59–80). John Benjamins Publishing Company.

Sood, S. O., Churchill, E. F., & Antin, J. (2012). Automatic identification of personal insults on social news sites. *Journal of the American Society for Information Science and Technology*, 63(2), 270–285.

Spencer-Oatey, H. (2008). Face, (Im)politeness and Rapport. In H. Spencer-Oatey (Ed.), *Culturally Speaking: Culture Communication and Politeness Theory* (pp. 11-47). Continuum.

Sperber, D., & Wilson, D. (1995). Relevance: Communication and cognition (2nd Edition). Blackwell. (Original work published in 1986).

Sproull, L., & Kiesler, S. (1991). *Connections: New ways of working in the networked organization.* MIT Press.

Stapler, J. (1995). The Expression of Disagreement. In K. Ehlich, J. Wagner, & D. Gruyter (Eds.), *The Discourse of Business Negotiation* (pp. 275–288). Mouton de Gruyter.

Statista Research Department. (2019). *Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019 (in millions)*. [Photograph]. Statista.

https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

Sterner, G., & Felmlee, D. (2017). The social networks of cyberbullying on Twitter. *International Journal of Technoethics (IJT)*, 8(2), 1–15.

Stivers, T. (2013). Sequence Organization. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 191–209). Wiley-Blackwell.

Stommel, W., & Koole, T. (2010). The online support group as a community: A micro-analysis of the interaction with a new member. *Discourse Studies*, 12(3), 357–378.

Suroiya, H. M. (2017). *Various verbal disagreeing strategies applied by male and female villains in Detective Conan movie series* [Doctoral dissertation, University Islam Negeri Sunan Ampel Surabaya). http://digilib.uinsby.ac.id/17955/

Tagg, C. (2015). *Exploring digital communication: Language in action*. Routledge.

Tagliamonte, S. A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Wiley-Blackwell.

Tannen, D. (2002). Agonism in academic discourse. *Journal of Pragmatics*, 34, 1651–1669.

Tannen, D., & Kakavá, C. (1992). Power and solidarity in Modern Greek conversation: Disagreeing to agree. *Journal of Modern Greek Studies*, 10(1), 11–34.

Taylor, C. (2015). Beyond Sarcasm: The Metalanguage and Structures of Mock Politeness. *Journal of Pragmatics*, 87, 127-141.

Terkourafi, M. (2005). Beyond the Micro-level in Politeness Research. *Journal of Politeness Research*, 1(2), 237–262.

Thatfunkymotherfunker. (2010). *Go fuck yourself*. Urban Dictionary. Retrieved 22nd February 2022, from https://www.urbandictionary.com/define.php?term=Go%20fuck%20yourself

Thurlow, C., Lengel, L., & Tomic, A. (2004). *Computer mediated communication*. Sage.

Tiwana, M. S., Javeed, F., Lali, M. I., Dar, H., & Bilal, M. (2018). Comparative analysis of context based classification of twitter. *2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA)* (pp. 1–5). Institute of Electrical and Electronics Engineers.

Toomaneejinda, A., & Harding, L. (2018). Disagreement practices in ELF academic group discussion: verbal, nonverbal and interactional strategies. *Journal of English as a Lingua Franca*, 7(2), 307–332.

Tracy, K, & Hodge, D. M. (2019). Discourse features of disputing in small claims hearings. In M. Evans, L. Jeffries, & J. O'Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (pp. 271-285). Routledge.

Twitter—About Verified Accounts. (2022). Retrieved 22nd February 2022, from https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts

Twitter—Developer terms. (2022). Retrieved 22nd February 2022, from https://developer.twitter.com/en/developer-terms/agreement

Twitter—Hateful conduct policy. (2022). Retrieved 22nd February 2022, from https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy

Twitter—How to report content and assign moderators. (2022). Retrieved 22nd March 2022, from https://help.twitter.com/en/using-twitter/report-content

Twitter—How to retweet. (2022). Retrieved 22nd February 2022, from https://help.twitter.com/en/using-twitter/how-to-retweet

Twitter—How to tweet. (2022). Retrieved 22nd February 2022, from https://help.twitter.com/en/using-twitter/how-to-tweet

Twitter—Our range of enforcement options. (2022). Retrieved 22nd February 2022, from https://help.twitter.com/en/rules-and-policies/enforcement-options

Twitter—Sensitive media policy. (2022). Retrieved 22nd February 2022, from https://help.twitter.com/en/rules-and-policies/media-policy

Twitter—Twitter privacy policy. (2022). Retrieved 22nd February 2022, from https://twitter.com/en/privacy

Twitter—User Agreement. (2022). Retrieved 22nd February 2022, from https://twitter.com/en/tos

Twitter—Violent threats policy. (2022). Retrieved 22nd February 2022, from https://help.twitter.com/en/rules-and-policies/violent-threats-glorification

United Union. (2020). *United Nations Strategy and Plan of Action on Hate Speech: detailed guidance on implementation for United Nations field presences*. https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20PoA%20on%20Hate%20Speech_Guidance%20on%20Addressing%20in%20field.pdf

Van Dijk, T. A. (2003). Critical Discourse Analysis. In D. Schiffrin, D. Tanne, & H. E. Hamilton (Eds.), *The Handbook of Discourse Analysis* (pp. 352–371). Blackwell.

Visansirikul, V., & Kitisin, S. (2018). Identifying Influencers with Ensemble Classification Approach on Twitter. *2018 22nd International Computer Science and Engineering Conference (ICSEC)* (pp. 1–4). Institute of Electrical and Electronics Engineers. http://doi.org/ 10.1109/ICSEC.2018.8712720

Vuchinich, S. (1990). The sequential organisation of closings in verbal family conflict. In A. Grimshaw (Ed.), *Conflict talk: sociolinguistic investigations of arguments in conversations* (pp. 118–130). Cambridge University Press.

Wallace, P. (1999). *The Psychology of the Internet*. Cambridge University Press.

Walther, J. (1996). Computer-mediated communication: Impersonal, interpersonal and hyperpersonal interaction. *Communication Research*, 23(1), 3–43.

Wang, W., Hernandez, I., Newman, D. A., He, J., & Bian, J. (2016). Twitter analysis: Studying US weekly trends in work stress and emotion. *Applied Psychology*, 65(2), 355–378.

Warne, R. T. (2018). *Statistics for the social sciences: A general linear model approach*. Cambridge University Press.

Waseem, Z., Davidson, T., Warmsley, D., & Weber, I. (2018). *Understanding abuse: a typology of abusive language detection subtasks*. arXiv. https://arxiv.org/abs/1705.09899

Waseem, Z., & Hovy, D. (2016). Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. *Proceedings of NAACL-HLT 2016* (pp. 88–93). Association for Computational Linguistics. https://aclanthology.org/N16-2013.pdf

Watanabe, H., Bouazizi, M., & Ohtsuki, T. (2018). Hate speech on twitter: A pragmatic approach to collect hateful and offensive expressions and perform hate speech detection. *IEEE Access* (pp. 13825–13835). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/ACCESS.2018.2806394

Waters, S. (2012). "It's Rude to VP": The Cultural Semantics of Rudeness. *Journal of Pragmatics*, 44(9), 1051-1062.

Watts, R. J. (1989). Relevance and relational work: linguistic politeness as politic behaviour. *Multilingua*, 8, 131-166.

Watts, R. J. (2003). *Politeness*. Cambridge University Press.

Watts, R. J. (2005). Linguistic Politeness and Politic Verbal Behaviour: Reconsidering Claims for Universality. In R. J. Watts, S. Ide, & K. Ehlich (Eds.), *Politeness in*

Language: Studies in its History, Theory and Practice, (pp. 43–69). Mouton de Gruyter. (Original work published in 1992).

Weisser, M. (2020). Speech acts in corpus pragmatics: Making the case for an extended taxonomy. *International Journal of Corpus Linguistics*, 25(4), 400–425.

Werkhofer, K. (1992). Traditional and modern review: The social constitution and the power of politeness. In R. J. Watts, S. Ide, & K. Ehlich (Eds), *Politeness in language: Studies in its history, theory and practice* (pp. 155–197). De Gruyter.

Wierzbicka, A. (1985). Different Cultures, Different Languages, Different Speech Acts. *Journal of Pragmatics*, 9, 145–161.

Wierzbicka, A. (2011). Bilingualism and cognition: the perspective from semantics. In V. B. Bassetti (Ed.), *Language and Bilingual Cognition* (pp. 191–218). Psychology Press.

Williams, A. M. (2005). Fighting words and challenging expectations: Language alternation and social roles in a family dispute. *Journal of Pragmatics*, 37(3), 317–328.

Wulczyn, E., Thain, N., & Dixon, L. (2017). Ex machina: Personal attacks seen at scale. *WWW '17: Proceedings of the 26th International Conference on World Wide We*b (pp. 1391–1399). International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/3038912.3052591

Xie, K., Miller, N. C., & Allison, J. R. (2013). Toward a social conflict evolution model: Examining the adverse power of conflictual social interaction in online learning. *Computers & Education*, 63, 404–415.

Yang, P. (2010). Managing mianzi in Mandarin Chinese Talk-in-Interaction: A Nonverbal Perspective. *Semiotica*, 181(1–4), 179–223.

Yang, Y. (2013). Gender differences in pragmatic strategies of disagreement in Chinese. *Proceedings of the International Academic Workshop on Social Science,* 50(1), 929–934.

Yan, W., Zhang, Y., & Bromfield, W. (2018). Analyzing the follower–followee ratio to determine user characteristics and institutional participation differences among research universities on ResearchGate. *Scientometrics,* 115(1), 299–316.

Yuan, X., Schuchard, R. J., & Crooks, A. T. (2019). Examining emergent communities and social bots within the polarized online vaccination debate in Twitter. *Social media+ society*, 5(3).

Yu, D., Xu, D., Wang, D., & Ni, Z. (2019). Hierarchical topic modeling of Twitter data for online analytical processing. *IEEE Access* (pp. 12373–12385). Institute of Electrical and Electronics Engineers. http://doi.org/ 10.1109/ACCESS.2019.2891902

Zhang, W., & Kramarae, C. (2014). 'SlutWalk' on connected screens: Multiple framings of a social media discussion. *Journal of Pragmatics*, 73, 66–81.

Zhang, Z., Robinson, D., & 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, & M. Alam (Eds.), *The Semantic Web. ESWC 2018. Lecture Notes in Computer Science* (pp. 746–760). Springer.

Zhao, X. X. (2018). *Emoji Use in Social Media* [Unpublished master thesis]. University College Dublin.

Zhu, H. (2008). Duelling languages, duelling values: Codeswitching in bilingual intergenerational conflict talk in diasporic families. *Journal of Pragmatics*, 40(10), 1799–1816.

Zubiaga, A., Liakata, M., Procter, R., Hoi, G. W. S., & Tolmie, P. (2016). *Analysing how people orient to and spread rumours in social media by looking at conversational threads*. arXiv. https://arxiv.org/abs/1511.07487