

*4

An overview of responses to (perceived) trolling

Claire Hardaker¹

Abstract

Computer-mediated communication (CMC) provides many benefits, including quick, efficient communication over time and space. At the same time, however, the anonymity it offers can give a sense of impunity, an illusion that behaviour is less hurtful than it really is, and a suppression of empathy. In short, CMC can be a fertile ground for conflict, and one particular manifestation of this is trolling. Trolling involves deliberately attacking others online, typically for amusement's sake. In some cases, it can be taken to such an extreme that it clearly violates UK legislation on hate-speech, abuse and menace. Whilst forensic linguistic research into threatening and abusive language is, however, gradually growing (Carney, 2014; Chakraborti, 2010: 99–123; and Fraser, 1998), there is a shortage of research into linguistic aggression online, and particularly research into trolling (see, however, Binns, 2011; Herring *et al.*, 2002; and Shin, 2008). In endeavouring to contribute to this under-researched area, this paper seeks to address the question, 'How do users respond to (perceived) trolling?' The answer to this is elaborated through the creation of a working taxonomy of response types, drawn from 3,727 examples of user discussions and accusations of trolling which were extracted from an eighty-six million word Usenet corpus. I conclude this paper by discussing the limitations and applications of this research.

Keywords: Aggression, computer-mediated communication, conflict, deception, impoliteness, Internet, manipulation, troll, trolling.

1. Introduction: under the bridge

Extensive research into threat, risk and conflict can be found in many fields beyond linguistics, including politics, business, sociology, law, psychology and peace studies. Within linguistics, however, research into what we might loosely call forensic pragmatics, such as speech acts of threatening, the use of impoliteness, the issuing of insults, and so forth, is far less established (see, however, Carney, 2014; Chakraborti, 2010: 99–123; and Fraser,

¹ Lancaster University, County South, Lancaster, LA1 4YL, UK
Correspondence to: Claire Hardaker, *e-mail*: c.hardaker@lancaster.ac.uk

1998). To complicate matters further, much of this research focusses on either speech, taken in some cases from contrived scenarios such as reality television shows, or from ‘traditional’ writing such as blackmail notes, threatening letters and newspapers. Linguistic research into computer-mediated communication (CMC) is only just beginning to gather momentum, and the combination of these factors presents a challenge. The intersection of linguistic research that examines both forensic pragmatic behaviours such as impoliteness, linguistic aggression, or threat on the one hand, and CMC on the other, is remarkably small. Smaller still is the subset of CMC research that looks at the specific type of online aggressive behaviour known as trolling (see, however, Binns, 2011; Herring *et al.*, 2002; and Shin, 2008).

Trolling is generally used to describe online antagonism undertaken for amusement’s sake (Hardaker, 2013: 77) – though, particularly within media and social networking circles, it is possible to find widely divergent denotations and usages that make the creation of any clear definition almost impossible. Current semantic complexities aside, trolling is a term with a history that can be traced back at least to the 1980s (e.g., Doyle, 1989; Maddox, 1989; and Mauney, 1982).² Despite the longevity of this behaviour’s existence – at least relative to the usual lifespans enjoyed within social networking – trolling has really only reached mainstream consciousness in recent years (see, for example, BBC, 2010; Camber and Neville, 2011; and Morris, 2011). And, as mentioned above, it has received little academic attention thus far, especially in fields like linguistics (see, however, Binns, 2011; Donath, 1999; Golder and Donath, 2004; Herring *et al.*, 2002; and Shachaf and Hara, 2010).

This great shortage of research has made it necessary to provide the groundwork to this field beginning with questions that are elementary, such as, “What is trolling?” and “How is trolling carried out?” Previous research (see Hardaker, 2010, 2013) demonstrated that current work on impoliteness largely fails to account for trolling, it set about establishing a working definition of this behaviour, and then outlined a range of major strategies said to be employed by trolls, including digression, (hypo)criticism, antipathy, endangerment, shock and aggression. This paper now takes another small step forward by considering how users respond to (perceived) trolling. To do so, it presents the findings of an analysis of 3,727 user discussions of trolling, drawn from an eighty-six million word Usenet corpus.

There are several points to note here, however, before moving on. For reasons that are given below, I focus purely on the perception of trolling, and I try to reiterate this throughout by using such terms as ‘(perceived) strategies’ and ‘(alleged) troller’. Where *perceived*, *alleged*, and so forth are omitted due to constraints of grammar or clarity, their

² Like Donath (1999), I distinguish between the ‘post’ (a troll), the ‘individual’ (a troller) and the ‘act’ (trolling).

inclusion should be assumed. Furthermore, *speaker* and *hearer* throughout should, unless stated otherwise, be taken to mean any producer of communication (e.g., one who speaks, writes and gestures) and any consumer of that communication (e.g., one who hears, reads and watches). This is an imperfect gloss since this data is literally written, but given that it also incorporates a wide range of speech-like elements, there is no ideal set of titles to use here. Moreover, given the frequency of *speaker* and *hearer* throughout, I have abbreviated them to S and H, respectively. And, finally, within much of the data, S's and H's sex are unknown or uncertain, leading to another complication. The use of the singular *they* is both awkward, and in some cases, confusing. Therefore, for purely alliterative convenience, S is deictically indexed as *shelher/hers*, whilst H is indexed as *hel/him/his* (e.g., 'S may find that, though she can prove H wrong, he rejects her explanation in favour of his own'). In the few instances where the examples already incorporate suitably identifying pronouns, these are followed instead.

1.1 The Gyges effect

CMC involves communication between humans that is mediated through a device such as a smartphone, computer or games console (Herring, 2003: 612). Many aspects of CMC are worthy of attention – anonymity, disinhibition, dehumanisation, the reduced ability to interpret intentions, and so forth – but due to the limitations of space, few can be considered. Perhaps the most crucial aspect is anonymity, since it underpins many anti-social online behaviours.

The worldwide web presents many examples of sites that can be shown on a cline from those that allow fully anonymous accounts, through to those that prohibit 'fake' identities and require full disclosure of offline identities (Bernstein *et al.*, 2011; Facebook, 2010; and Lampe and Resnick, 2004). The notion of a 'real' identity is problematic for reasons too numerous to tackle here; however, the practical reality is that users will typically circumvent or ignore software, administrators and terms of use, and choose their own level of online identity disclosure.

More relevant for this paper is the Gyges effect or, in other words, the ways that anonymity can affect behaviour. This has been of interest to academics across many fields, including psychology, sociology and philosophy, back to the time of Plato and his story of the shepherd, Gyges. This shepherd, upon finding a ring that makes him invisible (i.e., unidentifiable and, therefore, anonymous), used the protection that the invisibility afforded him to infiltrate the royal household, seduce the queen, assassinate the king and take the kingdom. Plato then argues that this power presents such a corruptive force that even the most morally upstanding could not resist it (Plato, 2007: 2.360b).

CMC can be viewed as a modern ring of Gyges which, through the invisibility or anonymity it offers, encourages a loss of self-awareness, a sense of impunity, an increased likelihood of acting upon normally inhibited impulses, increased polarisation, and decreased consideration and empathy for others online (Sia *et al.*, 2002; and Siegel *et al.*, 1986) – or, in short, a sense that users can carry out, and then hide from the consequences of anti-social and criminal online behaviour. This toxic disinhibition (Vinagre, 2008: 321), according to Douglas and McGarty (2001: 399), can manifest itself in behaviours such as trolling and flaming.³

A final, though by no means minor issue, is the way in which we deal with deceptive linguistic practices like trolling, since issues such as intention and interpretation become pressing considerations. Since these were discussed at length in Hardaker (2010) and again in Hardaker (2013), due to considerations of wordcount, they are only summarised here. In brief, following several researchers (e.g., Arundale, 2008; Gibbs, 2001; and Haugh, 2008), I take the view that we do not ‘know’ or ‘retrieve’ interpretations or intentions. Instead, it is my view that H must hypothesise from the available evidence, sometimes very quickly, what S actually intended (Culpeper *et al.*, 2003: 1552; and Mills, 2003: 136). This means that, as interactants, we are continually, recursively hypothesising about:

- (1) The intentions of others;
 - (2) *Their* interpretations of *our* intentions;
 - (3) *Their* hypotheses about *our* interpretations of *their* intentions;
- and so on.

(Mills, 2003: 45; and Mooney, 2004: 900)

Under normal circumstances, these hypotheses may well be unconscious, and only consciously called into question when we experience interactional problems. Furthermore, these hypotheses may be based on available contextual and cognitive information such as logic, schema, inferencing, and shared historical, cultural and social knowledge. However, because this process is still all largely educated guesswork, S or H can be (or can appear to be) incorrect for a number of reasons, including mistakes, ambiguity, deception, and so forth (Grimshaw, 1990: 281). This grey area, however, is precisely the opening that allows trolling to exist. Since we do not ‘know’ another’s intentions, an indeterminate area is available for the troller to exploit – an ‘unknown zone’ in which they can imply sincerity, whilst intending mischief; and since it is impossible ever to ‘prove’ what was really intended, there is always room for doubt and disbelief from H, and for deception, insincerity and manipulation from S.

A further result of this ‘unknown zone’ is that we – as analyst or user or bystander – cannot ‘know’ a troller’s guilt, and that means,

³ It is beyond the scope of this work to compare flaming and trolling; however, flaming should be understood as a sincere (over-)reaction to provocation.

furthermore, that there is simply no method in existence by which we can retrieve ‘every example of trolling’, for instance. First, a strange irony of trolling research is that the most skilled, covert types of trolls will never be identified as such. They will always successfully evade or defend themselves against accusations of trolling, and will, therefore, appear to be just like any other sincere member of a CMC group. Secondly, as discussed above, to claim to ‘know’ that someone is trolling is to claim to ‘know’ their intentions, and yet this is not possible. We can guess at, hypothesise, reconstruct, and co-construct, and so on, but we cannot ‘know’. And, thirdly, this would ignore the fact that the identity of trolling is often co-constructed, sometimes through heated arguments spanning many days, involving dozens of users, and situated within a set of fluctuating, community-based norms. As a result, extracting instances of trolling from the data is not as straightforward as it might appear, and this is discussed next.

2. Data and method: troll hunting

The data in question comprises two corpora derived from Usenet (described below) with a combined word-count of 86,412,727 words.⁴ Usenet predates the current incarnation of the worldwide web, and is:

[...] an electronic forum for discussion of almost any subject, allowing access to millions of computer users who share similar (or very different) hobbies, interests and worldviews (McLaughlin, Osborne, and Smith, 1995). Characterised by its immediacy and sheer volume of traffic, Usenet groups based around the discussion of a particular topic afford a prime example of Internet communities. The main method of communication is text-based e-mail, although some groups permit the exchange of graphics, sound or video files.

(Baker, 2001: 1)

Datasets from Usenet have several benefits: (i) newsgroups exist on an extraordinary range of topics; (ii) some have archives spanning back to the 1980s; and (iii) posts can be downloaded. The result is that building a principled corpus with regards to topic, chronology, language, region, and so forth is relatively easy. A further aspect is that trolling is said to have begun on Usenet (Tepper, 1997) and, indeed, we can find examples of this term’s usage in archives up to three decades ago (e.g., Maddox, 1989; Mauney, 1982; and Miller, 1990). However, as with any other type of data, this type of data has its limitations, too. Usenet is arguably a dying platform – a fate perhaps accelerated by the rise of Web 2.0 social networks. As a

⁴ See: <http://groups.google.com/group/rec.equestrian> and <http://groups.google.com/group/uk.sport.football>

result of this, the average user demographic *may* tend towards older long-term members, with the result that behaviours found on Usenet may be different from those found on Twitter, Facebook, *etc.* This is nearly impossible to rigorously test, however; and, overall, for the purposes of this exploratory article, the benefits outweigh the drawbacks.

The first corpus, RE, was built from a subset of the rec.equestrian newsgroup. RE's theme is equestrianism, competitions, breeding and related topics such as welfare, agriculture and nutrition. The second corpus, SF, was created from a subset of the uk.sport.football newsgroup. SF's theme is (English) football, fixtures, leagues and related topics such as wages, management and refereeing. Both groups are unmoderated, and allow users to employ high levels of anonymity. Further details of each corpus are found in Table 1.

==Insert Table 1 about here==

Once RE and SF were created, I used WordSmith (Scott, 2009) to retrieve all instances of TROLL*.⁵ Searching the corpora with an open-ended wildcard resulted in around ~9 percent false hits (e.g., *Trollope*); but using this wildcard also retrieved derivations, inflections, compounds, neologisms, and some typographical errors that might otherwise have been excluded. RE returned 2,643 instances, whilst SF returned 1,456 instances. This created an initial sub-corpus of 4,099 examples that was reduced to 3,727 once the false hits were excluded. Though WordSmith retrieved an impressive set of results from RE/SF, no search can currently retrieve off-record or implicit references to TROLL (e.g., 'it has a sub-bridge apartment'). These instances were only captured if TROLL occurred in a more explicit part of the thread. The examples discussed in this paper should not, therefore, be taken as an exhaustive collection of all instances of trolling in RE/SF.

A cautionary note, as already mentioned above, is that we cannot know someone's trolling guilt or innocence; so, should a genuine troller successfully encourage many users to respond without anyone ever realising what she is about, then such an example will not have been captured by this search. This is because the search looked purely for TROLL* and if this string or any of its variants did not occur in the thread, then the thread was not retrieved.

This brings to the fore an issue in combining corpus linguistic methods with pragmatic analyses. As a field, pragmatics concerns itself with the constructions and understandings of meanings within social interaction, including implied and inferred meanings, intentional and unintentional meanings, the dynamic and the emergent, or, in short, with phenomena that may leave little or no trace on the text whatsoever

⁵ The asterisk (*) wildcard denotes 'zero or more characters', so a search for CAT* retrieves *cat, cats, catch, etc.*

(Culpeper and Hardaker, forthcoming). Given that corpus linguistics relies heavily on stable linguistic features that can be consistently found and readily counted, it would be easy to assume that using corpus methods to carry out pragmatic analyses is simply not viable. However, in reality, the field of ‘corpus pragmatics’ is gathering momentum, as evidenced by the publication of major volumes and overview papers (e.g., Jucker, 2013; Jucker *et al.*, 2009, Romero-Trillo, 2008; and Taavitsainen *et al.*, 2014).

In the case of this study, because it is not possible to retrieve pragmatic phenomena such as trolling by searching for fixed linguistic forms, it was necessary to start more simply, by investigating the metapragmatic label of *trolling* first and working outward to analyses of (alleged) trolling, thereafter. This meant that the computer’s role in the analysis largely ended after retrieving the varying linguistic forms (*troll*, *trolls*, *trolled*, etc.), since, at this stage, the aim was then to identify the circumstances that had led to a version of the meta-pragmatic label occurring, and then, for the purposes of this paper, to identify from those examples, the major types of response to (alleged) instances of trolling. In short, in this case, a corpus approach was useful in rigorously deriving from the full corpus a smaller dataset of examples that were then amenable to manual, qualitative analysis; but the extent of corpus methods within this forensic pragmatic analysis should not be overstated. (This issue is discussed further in Section 2.2.)

2.1 Classifying responses

Following Watts (2003: 9), the analysis of first-order understandings and co-constructions of interaction is viewed as the only valid means of developing a social theory – not just of politeness, but also of impoliteness, trolling, and so forth. This is because first-order approaches foreground the interactant’s interpretations, whilst trying to background the analyst’s. This approach is not problem-free, however. Lay users often lack the meta-linguistic repertoire required to apply a rigorous framework to their interaction (O’Keefe, 1989). Despite this, interactants can and do regularly assess behaviours such as trolling on a moment-by-moment basis, and individuals often also arrive at interpretations that are consistent with others. This suggests, therefore, the existence of an implicit, principled system of assessing the contextual appropriacy of behaviour.

In-keeping with this first-order approach, at first, the subsequent classifications of responses were informed by whether they fell within the scope of the four strategies suggested by Harris *et al.* (1986), later developed by Culpeper *et al.* (2003: 1562–5), and then modified to account for trolling. These were:

- (1) Does not respond to the trolling;
- (2) Accepts the trolling;

- (3) Counters the trolling *defensively* (i.e., by protecting H's own face);
or,
- (4) Counters the trolling *offensively* (i.e., by attacking S's face).

As the analysis progressed, however, it quickly became apparent that these categories were largely unsuitable. For instance, it is virtually impossible to determine whether someone has chosen not to respond, and why that might be. It is also difficult to apply a concept like 'accept the trolling' to this data. Meanwhile, both offensive and defensive countering did occur, but there were still many other types of response that were left unaccounted for. As a result, this initial framework was altered and expanded to account for other response types. As far as possible, those categories were driven by the user's interpretations and discussions of their own responses surrounding the mention of trolling, (though inevitably, this was not always possible, since users do not always meta-label their own behaviour).

When categorising, it is possible to create potentially thousands of contextually specific, fine-grained sub-categories (e.g., responding sincerely with anger, responding sincerely with disinterest). However, the interest here was in identifying more general behaviours (e.g., responding sincerely), and this resulted in seven response types ranging from the fully naïve (i.e., apparently unaware of the trolling) through to the fully cynical (i.e., apparently fully aware of the trolling): engaging, ignoring, exposing, challenging, critiquing, mocking and reciprocating.

2.2 Counting responses

When we come to the issue of counting pragmatic strategies in any corpus, it is vital to emphasise two points. Firstly, given how creative users and groups can be, this list (and one might argue, any list of pragmatic strategies) cannot ever be truly called exhaustive. In addition, the data does not represent all of Usenet, nor even all equestrian or football groups on Usenet. Data from different sources, times and cultures will almost certainly reveal strategies that are not captured here. As such, these findings should be considered cautiously – as representative of these two groups over the timespan in question, and perhaps more importantly, as offering hypotheses to test against CMC groups. It would be interesting to see, for instance, if trolling is responded to in the same way in groups that are moderated, that encourage full identity disclosure or that only support short-form messaging.

Secondly, in keeping with the first-order approach, I strive to emphasise what the user emphasises. The issue here, however, is that different users may cite different response types for the same individual, or one user may cite multiple possible response types (e.g., that they are both exposing the troller and challenging her at the same time). And, of course, users do not always include that useful meta-discussion of the response

type they feel they are giving, thereby leaving the analyst in the position of ascribing one. Whilst this was not problematic in the majority of cases, it is always possible for misinterpretation to creep in. This returns us to the fundamental problem stated above: our inability to know, for certain, what another individual intends.

A third issue is that since all the examples were retrieved by searching for instances of TROLL*, and since, of these, nearly all involve Hs discussing their perspectives on trolling, any statistical analysis of the data would be biased towards features that Hs interpret as trolling, rather than those that definitely are instances of trolling.

The combination of these factors – (i) the potential co-occurrence of several types of response; (ii) the frequent absence of meta-discussion from H about which response type they feel they are using; and (iii) our general inability to know intention – has made statistical, quantitative analyses so highly qualified as to have little value in this study. This section presents, therefore, a qualitative overview of the major response types that could be identified when the data was analysed. In this case, ‘major response types’ is used to mean response types that: (i) occur consistently (e.g., three times or more in at least one dataset); (ii) always include at least one user voicing their suspicion that the post being responded to is a troll; and (iii) preferably occur across both datasets (though this was not obligatory).

2.3 Changes to the data

In line with previous papers (2010, 2013), as far as possible, I have avoided altering the data, and all spelling, grammar and (non-indenting) punctuation are original. However, some changes were necessary. These are as follows:

- (1) To enhance anonymity, all names are replaced with letters. In each example, user-letters re-start at A (though see point 7, below). This means that, unless stated otherwise, A in Example 1 is not the same person as A in Example 2.
- (2) Italics in square brackets indicate information that has been removed or glosses, (e.g., [*web address*]).
- (3) Indenting punctuation (e.g., >) has been removed, since subsequent replies already present chronological threads for the reader to follow.
- (4) For brevity and clarity, unnecessary line-breaks were removed.
- (5) Italics highlight the parts of an example that are being analysed.
- (6) Whilst Usenet now offers rich text, very few posts use this, so all are presented in standard text.
- (7) Where applicable, the (alleged) troller is designated A (underlined). Where none appears, A is avoided.

- (8) Most examples are excerpts from threads that are, in many cases, significantly longer. Inevitably, this leads to orphaned deictic markers, users being addressed but not appearing, missing contextual information, and so forth. To mitigate this, examples that stand reasonably well alone are used, and/or the relevant context is summarised.
- (9) Sections deleted from long posts are indicated using [...].

3. Analysis: user responses

This section considers how users respond to (alleged) trolls. It also addresses how well the replies fall within the scope of the response strategies suggested by Harris *et al.* (1986), developed by Culpeper *et al.* (2003: 1562–5), and modified here to account for trolling. As noted above, these involve: (i) not responding, (ii) accepting, (iii) countering defensively, or (iv) countering offensively.

3.1 Engage

The first response-type involved one or more users sincerely engaging with a troller, and, for the corpus software to have retrieved the thread in the first place, at least one other user explicitly communicating her suspicions that the post being responded to is a troll. Sincere engagement involved responses to the (alleged) troller's 'pseudo-intent' – that is, the deceitful, faux-sincere impression, given by the troller, of wishing to be a genuine group member – that suggested that those pseudo-intention(s) were genuine, and that any 'real' (i.e., trolling) intentions had not been detected. As in Example 1, sincere H-responses can include aggravated retaliation (i.e., flaming), interest, dismay and shock. These types of responses are not easily captured by the strategies outlined above by Culpeper *et al.* (2003: 1562–5), particularly if S's trolling strategy was off-record:

Example 1

- A** I spotted an old usable horse trailer, while taking my wife on an anniversary trip yesterday. So I drove out, and [*Name Surname*], the seller, pulled a tractor and backhoe out from in front of the trailer I wanted, so I could try a short tow, with my anemic S10 truck.
- B** You're going to try to pull a loaded horse trailer with an S10?
I think you're going to be minus an S10 in a hurry.
- C** He's a troll. You bit.

(RE060830)

In this case, B engages with A's apparently genuine intent by critiquing his choice of vehicle for pulling a horsebox, and, in so doing, constructs himself as someone providing a sincere answer to another user's sincere question. C, however, not only assesses A's behaviour as trolling, but also explicitly identifies that B has taken the bait, reframing B's response as a waste of time.

As seen in Example 1, sincere engagement with a troll is sometimes termed *biting*, *being trolled*, *being hooked*, etc. (note the derivations from fishing), and as with most instances where a person finds out that she has been taken in, there can be an associated loss of face or damage to pride. Experienced users endeavour to avoid being cast as one who has taken the bait due to the implications this has of naïveté, ignorance and gullibility, and they also strive to position themselves as knowledgeable members with the foresight to avoid, predict and deal with trolling. The greater the member's status, the greater the loss of face is if that member is then trolled:

Example 1

- A** Who is to say this fabled other recipient of Best's new liver would have treated it any better? What if it went to someone who had lived a decent, healthy lifestyle that had suddenly been struck down by liver disease because he had offered a smoker a lift to the fish factory one morning? What, then, if Mr Fabled Recipient then thought to himself, 'Fuck me, Nigel (for that is his name)! I have just wasted 42 years of my life being a stupid goody two shoes and yet I still nearly died of liver disease. Fuck this for a lark, I'm going on the biggest bender of all time.' And then our Nigel is found dead from snowball OD six months later. How can you say that wouldn't have happened? You know, George Best wasn't the only cunt on the planet capable of fucking up a new liver.
- B** How the fuck does offering (or even giving, for that matter) somebody a lift, to a fish factory, give nigel liver disease?
- C** **The question we all wanted to ask, but didn't want to feel trolled.** I love you, I do.

(SF060117)

Whilst B's response to the (alleged) troller, A, seems rather wry and sarcastic, it is C's response that most clearly encapsulates the tension between responding to someone out of curiosity, interest or irritation, and not wanting to expose oneself to the ignominy of being trolled. As Donath (1999: 6) states, 'trolling is a game about identity deception, albeit one that is played without the consent of most players'.

Sincerely engaging with a troller is akin to participating in a large, public game where one player is cheating, but, moreover, where other, honest players are aware of this. Whilst some honest players may try to educate the dupes, others may enjoy simply sitting back and mocking the

naïveté of those who have been fooled. This can, therefore, make being trolled a publicly embarrassing affair, and, as with impoliteness, the size of audience can play a pivotal role in the extent of face-damage that is involved.

3.2 Ignore

A strategy that occurred extensively in RE, but never (properly) in SF, was to ignore trolling through blocking, killfiling (i.e., blocking) or simply not reading the troller's posts. Where users do this silently (that is to say, without alerting others), there is, of course, no trace in the data, and this is most consistent with the 'do not respond' option suggested in Culpeper *et al.* (2003: 1562). However, this leaves nothing to analyse, nor even any evidence of its occurrence, so it is not discussed any further.

Instead, the focus is on the strategy of 'overtly' ignoring the troller, akin to Culpeper *et al.*'s (2003: 1562) defensive strategy of 'opting out on record'. Users commonly (purport to) act in the group's interest by alerting others to their suspicions, and suggesting methods of dealing with the problem. This strategy is predicated on the idea that trolls are seeking attention, and if sufficiently starved of it, they will eventually get bored and leave. Users, particularly in RE, frequently tried to curb responses from others by not only highlighting the suspect's status as a troller, but also by advising, asking or telling other members what to do:

Example 3

B Killfiles, girls. Use your killfiles and send this troll to the muck pile.

(RE030929)

Example 4

C FOLKS THIS IS A TROLL STOP FEEDING IT

(RE060515)

Example 5

D All I asked was "where have you been!?". I wasn't encouraging anything.

E Yes you were. **Trolls should be ignored, not engaged in conversation.**

(RE060807)

All three examples demonstrate the notion that killfiling or ignoring the troller in some way denies her what she wants. As such, B, C and especially E construct their identities as experienced members who not only know how to deal with trolls, but are willing to take steps to protect their group from disruption. Doing so positions trolling as an inappropriate,

undesirable behaviour that is socially marked in RE, and that needs to be stopped quickly to maintain group harmony. As mentioned above, however, this strategy was notable by its absence from SF, with one interesting exception:

Example 6

B **Ok, ok, ok, folks, let's not feed the troll and maybe he'll go away.** I've already warned him privately that censure will be taken if he continues to spam. There's only one newsgroup here for which this post is on-topic.

C **Oh well that's sure to scare him off!**

B lol It did, well, nearly.

(SF050417)

This is the only SF example containing a user commenting on a discussion about ignoring trolls. On further investigation, whilst C is a member of SF, B is a member of another group that has also received the (alleged) troller's post. When replying, B has then (probably accidentally) cross-posted back to all the groups that were caught up in the initial offending post, including SF. SF member C is more concerned with mocking B's attempts at solving the problem than supporting C's strategy of ignoring the troller. C, therefore, does not cast trolling as problematic or marked – rather, he constructs B's efforts as laughably futile.

Whilst RE users took frequent steps, sometimes heatedly, to prevent one another from responding to trolls, SF produced no examples of this type. In fact, the general trend, as seen below, is that SF treated trolling as a game (albeit sometimes a very hotly contested one). In RE, trolling was predominantly treated as aberrant and undesirable behaviour that required quick suppression. However, attempts to manage trolling sometimes caused more irritation than the trolling itself, since those efforts prolonged and drew attention to the unwelcome interaction:

Example 7

A Back to smells: ammonia doesn't work. [*Name*] like that so much he tried to put his nose in the baggie and snort up all the ammonia soaked cotton balls.

B Yep, take a plastic bag, fill it with ammonia soaked cotton balls, and put it over your horse's muzzle..... Real f**king smart.....

A Smarter than you, who appears to be really fucking DUMB.

C **PLEASE DON'T FEED THE TROLL!!!! damnit.**

B **It's my thread, I get to feed whomever I want in it. BTW, do you think posting "PLEASE DON'T FEED THE TROLL!!!!" a zillion times ISN'T feeding the troll? Silly you.**

F Fine. Whatever. Not feeding trolls is usually a choice of wisdom. Clearly some of you prefer to be taken in.

(RE060717)

In Example 7, C's efforts to prevent B from continuing to engage with A only aggravate B, despite the fact that B seems to accept the probability that A is trolling.

A notable feature of the ignore strategy was that in most examples, users choose to speak to one another, instead of addressing the (alleged) troller directly. However, indirect conversations between members which aim to exclude and stifle the troller, can still provide an opening to respond to, thereby actually exacerbating the problem. As B points out above, even an overt 'non-response' (e.g., where members openly declare their intentions to ignore the troller) is still a response, and automatically self-cancels its intended aim, since the troller can still reply to it.

3.3 Expose

Group members did not always have the choice of simply killfiling users that they wished to ignore, since sometimes, as noted by Donath's (1999: 45) work, and later extended by Utz (2005: 50), one strategy of (alleged) trolls is to knowingly present authentic-sounding but poor or dangerous advice. In the case of this dataset, and particularly RE, this also extended to giving out poor-quality information and setting dangerous examples. As a result, more knowledgeable members with social and moral concerns for other less-knowledgeable group members are effectively driven to contradict this advice. Some users further felt obliged to take note of those who had given bad advice previously, and refer back to this as a form of character reference:

Example 8

A Hi **B** – pretty – and they grow so fast.... My eye caught up on a detail of Mom's forefoot: I compared it with Mystic's forefoot in a shot on a trailing lead I took earlier this evening: [*link to photograph of a young horse grazing loose (i.e. not held by anyone), its reins thrown over a rock and its leadrope trailing on the ground around its forefoot*] ...and I noticed a difference in angle – just a passing step probably?

C There are so many scary things in your picture, **A**. I'm not going to rise to the troll's bait. In re your observation about angles: Joints look different depending on whether they are standing still or moving. That's how joints work. Duh.

A Sorry if I frightened you, but I'm glad you spoke up. **I already read what happens when someone here tries to do the kind of simple thing I do without a problem**, I need a cautionary note every time.

D That's because there's a huge difference between safe practices, and **unsafe practices where you have been lucky, so far. [...]**

Given your propensity for doing unsafe things, some day your luck will run out too.

- C** Some of the time you will get away with leaving a young horse tacked up and grazing with a rope on the ground. But just like when you pass on the blind corner, when you don't get away with it, the result is so messy, and often so tragic, that people with functioning brain cells find that taking reasonable precautions (AKA good horsemanship or good driving) is smart. Failure to take these precautions is the origin of most Darwin Awards. Sheesh.
- A** Hi C. No offence intended. But may I ask, are you a rated riding instructor, or possibly a show judge? I have found it's very important to check my sources when reading this NG. And I accept some risks because I am handling several green horses every day – nearly always quite alone. The picture that got you excited was a snap I took after ground walking two young geldings in tack yesterday. I dropped both their lines while I took a picture. I walked another two today. And I dropped their lines for a few moments too. And probably will continue to do so every day. Sorry.
- B** No, I'm not a "rated" riding instructor, whatever that is. **I do, however, know better than to leave a fully tacked up young horse loose with a long rope while I take pictures to impress rec.eq.**

(RE070201)

Example 9

- A** Rendered animal fat – such as neatsfoot oil for leather tack – is probably as close as animal fats should get to horses I think – if that. I add a splash of Canola to feed, along with sunflower seed.
- E** **And I doubt I will *ever* follow any of your advice, troll.**

(RE090125)

In this instance, A in both examples is the same individual, and despite the fact that the advice they offer in

Example 9 is not momentous or particularly risky, E is quick to discount it based on A's behaviour such as that in Example 8 from two years previously. As a result, though E may feel morally unable to killfile A due to his habit of giving out bad advice that needs to be contradicted, she can still informally moderate his interaction in other ways, by drawing attention to his status as an untrustworthy outcast.

These two examples also raise an important issue about assigning roles to users. The 'defensive' counter-strategy arguably most closely fits with C, D and E's responses, yet still falls far from the mark. In both examples, C, D and E are not really protecting their own faces – rather,

they are trying to protect other, innocent users from potential harm, and in so doing, they are arguably using ‘offensive’ strategies that damage A’s face. Further, if we accept that A is trolling, then he is using a covert offensive attack to trigger an overt (offensive or defensive) counter-attack which will then look like the prime aggressing action, rather than a (justified) reaction. In so doing, he attacks whilst assuming the position of the victim, and the targets (the group) are positioned as the attackers if they attempt to defend themselves. In short, even to categorise a post as an attack or counter-attack, we must first establish each participant’s role, since the proactive and wrongful aggressor may have assumed the identity of the reactive and wronged victim, whilst seeking to place the actual victim in the position of aggressor.

3.4 Challenge

Within the data, it was clear that not just any response was a success for trolls. Some users adopted responses that were offensive in character (Culpeper *et al.*, 2003: 1562–5; and Harris *et al.*, 1986), but that aimed to avoid the appearance of having been trolled, whilst also allowing themselves the opportunity of responding to (perceived) trolling. This occasionally occurred indirectly, where users would talk amongst themselves, usually about how trolls ought to be dealt with:

Example 10

B Haven’t we used up all the troll food yet?

C **Yeah, but I’ve got some D-Con left. And a night scope on my shotgun.**

(RE031201)

Example 11

D Personally, I’m thinking the most appropriate treatment would be a 150 grain copper/lead bolus administered intracranially, but unfortunately, that’s considered an "off-label" use for trolls, and there’s so much paperwork involved in getting approval that it isn’t worth the effort.

(RE060622)

The benefit of an indirect response is that, should a troller respond, the user can deny that his post was about her, but indirect challenges also carry the risk that the troller may not realise, or may choose to pretend not to realise, that these posts target her. Indeed, given the nature of this type of response, it may be considered as a hybrid of both exposing (see previous section) and challenging.

Challenging also occurred directly, where the user would openly confront the troller:

Example 12

A I am a civilized person. If you defeat me that is fine.

B Fine or not, you went down in flames when you entered this newsgroup. **Even the dullards on this bunch eat trolls breakfast and use their bones for toothpicks.**

(RE051108)

Such attempts, however, whilst they may seek to scare a troller off, can easily be interpreted as flaming, and the user may then find his response being reconstructed by the troller and/or other members as taking the bait and being trolled – in short, as sincere engagement, particularly if the troller is amused enough by his response to keep the interaction going. As a result, this is a strategy with a higher risk–reward balance. Should the user successfully drive the troller out of the group, he has the satisfaction of proactively enforcing his group norms (rather than simply ignoring the troller in the hope that she will get bored and leave) but if his responses only entertain the troller and cause greater disruption, he may find himself exposed to the ignominy of being censured by other members and being viewed as a novice who has been trolled.

3.5 Critique

One of the most interesting aspects that emerged was how trollers are appraised for their trolling quality, especially in terms of how others responded. The efforts of (alleged) trollers were open to critiques of their cleverness, effectiveness and success, and users from both RE and SF were quick to criticise ‘poor’ attempts:

Example 13

A I am boarding my pony at a stable and the dumb bitch that runs the place is having a fit because I am 3 months late in my board money. I told her I planned to pay in full as soon as I win the lottery this month, and I know I will be winning because it's my turn to win. Instead of being happy that I plan to get paid up, she mailed me a letter that said "You must pay all your back board in full by May 15, 2006, or you will be dealing with the po lice". (Exactly in those words). I know that po lice means pony lice. The bitch is going to intentionally infect my pony. That is not fair to the pony. I think I am going to have to sue this bitch. By the way, If anyone plans to board their horse in the [Location] area, DO NOT board at [Business Name]. The owner is a fucking bitch, and she beats and starves all the boarders horses.

B Well aren't you a friendly girl. **You might have to keep practising your trolling though. It's not very good.**

C Worst troll I've seen in a LONG time. Must be a preteen.
(RE060504)

Example 14

D Please make more of an effort when you want to troll.
(SF050419)

Example 15

E Bad troll, back in your box! I would be ashamed of crappy troll material like that...
(SF070420)

Perhaps unsurprisingly, A's efforts are roundly criticised, and, intriguingly, C, D and E suggest that trollers should actually try harder and come up with better material. This strategy is notable for two reasons. It (implicitly or explicitly) positions the users as having already identified the accused as a troller, and it also constructs the troller as poor, inept or unintelligent. By comparison, the users present themselves as knowledgeable, experienced with trolling, and in a position to advise and appraise the efforts of would-be trollers.

This strategy was also used to judge the (alleged) troller's entertainment value:

Example 16

B Why is anyone responding to this troll? In the history of the trolls we have had here, **he is not entertaining.**
C Yes, but he's a new troll. And he's really dumb, no matter what he says. **Don't tell me you've never watched a cat tossing around a dead mouse and been amused...**
(RE050907)

Example 17

A You know not who you are dealing with...
D Indeed I do- **an exceptionally inept and entertaining troll.**
(RE060112)

The efficacy of this strategy may be found in the fact that by turning trolling into a show or game, the users become an audience that has consciously chosen, and is allowing itself, to be amused. Meanwhile, the (alleged) troller loses the advantage of any supposed deception through which they might aggravate the group, and is instead turned into a caricature that is being judged for her ability to put on an amusing show. However, adept, clever and successful trollers were also often directly (see Examples 18 and 19) and indirectly (see Examples 20, 21 and 22) appraised and critiqued by users. Interestingly, praise for 'good' trolling was markedly less common in SF:

Example 18

A Ah...now it makes sense. You *are* a troll! I suspected as much. I give it to ya, **you were a bit (only a bit) more clever than most at the start** but you couldn't help yourself could you...

(RE060531)

Example 19

B Ok, I get it. **Nice subtle troll.** Fair play to you, you had me going there for a bit..

(RE060317)

Example 20

C **It was a successful troll.** I don't know why people do it. It must be a sort of perverted hobby, like graffiti or crapping in dressing rooms.

(RE070114)

Example 21

D One thing I can say for **E** is that **he's the most successful troll here**, and everyone keeps contributing to his cause! Duh. STOP already.

(RE070213)

Example 22

F **I think we are dealing with an exceptionally talented troll!** She even has us talking about her in another thread. Without feeding the troll, we are feeding the troll.

(RE060317)

This suggests that even though trolling is meant as an aggravation to users, it can become a two-sided game where a troller seeks to deceive and attack, and users parry with critiques on quality and cleverness. In so doing, users are addressing the (alleged) troller's real intent (i.e., to troll), rather than her pseudo-intent (i.e., to sincerely engage the group). Where this occurs, users appear to adopt something close to mock impoliteness. However, whilst mock impoliteness aims to enhance social cohesion and affect (and may, indeed, do so between the group members in question), in these cases, the troller is still ostracised and excluded (Haugh, 2008, 2010; Labov, 1972; and Leech, 1983: 144).

3.6 Mock

Users would occasionally offensively counter-attack a trolling attempt by mocking or parodying the troller. This occurred both indirectly in

discussions about trollers, and was sometimes aimed directly at the troller. Whilst it occurred occasionally in RE, it was most popular in SF:

Example 23

- A** Take your little net-kiddie wars to another dimension in time. If you are men, I feel embarrassed for all who reply -- because you act like women.
- B** **I refuse to respond to this obvious troll.**
- C** Errrr...didnt you just reply?
- D** You just did!

(SF060601)

Example 24

- E** geez, it's tough to be a troll on this group....<g>
- F** how come [Name], [Name], [Name], [Name], etal, haven't joined in? <snort>
- G** Too much pride to waste their talent on a whack job? ;-)
- H** Aw, the pickin's have been sorta thin in the troll department lately...you keep your hand in on what's available, ya know? ;>)
- J** **My horse is laying on the ground barely breathing. Until I can get medical/surgical on him and can then call the vet (not before next week), can you tell me: 1. what's wrong with him? and 2. do I even deserve a horse? Thanks in advance! Love, trollerita**

(RE051007)

B, a long-established user with a reputation for humour, uses a reply to a troller's efforts as an opportunity to not only potentially troll the troller but also to mock the troll, as a method of enhancing her solidarity with her own group. Similarly, J draws on in-group knowledge and norms to parody the types of trolling that are recognisable, and therefore are amusing to, her group (de Fina, 2006: 352).

Indirectly mocking a troller incorporates its own element of face-protection since the users could deny the relationship between their humour and the troller. This strategy also serves the function of strengthening group cohesion by identifying those in-group members who can understand and enjoy the humour appropriately from those who are excluded or out-grouped – either by a lack of knowledge or because the humour targets them.

Cultural differences between RE/SF may also be one reason why this strategy occurs more frequently in SF. Specifically, SF is characterised by a more jocular, relaxed and satirical nature, unlike RE which tends more towards serious discussion and is less tolerant of behaviour that falls outside the group's norms of interaction. As a result, SF users more frequently challenged the troller directly by using mockery or sarcasm:

Example 25

B Heh. I'm starting to like you, A, I really am. It's precisely your self-righteous attitude which is going to save the world from everything. You're my second-favourite troll of 2009. Although I do actually suspect you of being **C**.

(SF090123)

In this case, B works to neutralise the threat that the alleged troller, A, poses by reframing her and her efforts from being a potential danger to the group's harmony and cohesion, to an entertainment that can be enjoyed, particularly by B, but also by the rest of the group. By further classifying A as only his 'second' favourite troll (and only 'of 2009'), this adds the extra insult of suggesting that A is not even the best at this, either. In some instances, this tactic was taken into the realms of 'educating' the troller, and therefore had a great deal of overlap with the strategy of critiquing trolling (see previous section). In others, SF users employed multiple strategies, such as mocking the troller directly whilst also taking up the running joke between themselves. In so doing, rather than (explicitly) suppressing or preventing the (alleged) troller from responding, the users even encouraged further replies that allowed them to continue their mockery. The intriguing result of this was that, as threads progressed, some trollers increasingly failed to reply, perhaps as a result of discovering that they had gradually been positioned as the one being trolled. As a result, these types of response at times overlapped with, and in some cases gradually developed into reciprocal responses (see below).

3.7 Reciprocate

A final major response type found in these datasets involved users endeavouring to jeopardise the troller's success by trolling them in return:

Example 26

C Umm... **B**? Do you think A made up her hot_ail address all special for us, just for this post? (Google is your friend.) Do you think she really has a husband? Do you think she is really even a _she_? Wait. I get it!!! **You're trolling the troll**. Had me going for a minute there. <g>

(RE060318)

Example 27

A There's a delete key?

F Assuming you haven't deleted it, then yes, yes there is. It's right next to the 'ignore' key.

E **Good ol' F. Always trolling teh trolls. I do hope life is treating you well.**

F Oh it is.

G You know, he's just got a big shiny new van with the words 'Good ol' **F** - Always trolling teh trolls' printed in big red letters on the side. And his number so you can call him any time day or night when you want some trolls trolled. I think the business will do really well. It's also got a "How's my driving?" sticker on the back but cleverly the phone number goes to his own house rather than the mysterious driving police.

(SF060330)

Examples 26 and 27 demonstrate the ability of users such as C and F to entirely switch roles from target to attacker, by taking on a trolling identity and attempting to aggravate the (alleged) troller. As noted above, this strategy can be born out of users challenging, critiquing and mocking (alleged) would-be trolls. In some instances, this appeared to be successful in that the troller ceased to take part in the group's discussions. However, in other cases, the high level of risk inherent in this retaliation-in-kind strategy became extremely apparent. In some cases, it merely triggered extensive, heated conflict spirals that spanned many days, hundreds of posts and dozens of users (Andersson and Pearson, 1999; Felson, 1982: 245; and Lein and Brenneis, 1978: 301). In two notable instances, however, which have not been reproduced to protect all involved, this escalated beyond trolling, into far more serious behaviour that is better captured as cyberstalking. In these cases, the (alleged) trolls deliberately found contact information about the users who had attacked them, including names, addresses and contact details of their workplaces and family members (including children). This information was then published in posts (an action that is informally known as doxxing) with threats or claims that the (alleged) troller had sent out e-mails to the various contacts linking them back to the argument and threatening the target with the loss of their employment if they continued it. In short, whilst this response type was, for some users, an amusing and inconsequential strategy, for a minority of (alleged) trolls, it was sufficiently aggravating to provoke them into at least threatening, if not actually trying, to cause off-line harm or damage to the user.

It is, of course, impossible to know whether those efforts met with any kind of success – one can only hope not – but this returns us to the serious consideration that whilst many manifestations of trolling may fall into the category of merely annoying another, this behaviour can also escalate, and sometimes extremely quickly, through a broad grey area of behaviour that is border-line illegal (with reference to UK legislation), through to behaviour that is menacing, grossly offensive, and seeks to do actual harm, whether social, financial, psychological, and so on, to the target.

1. Conclusion: over the bridge

In this paper, I analysed the common responses to (alleged) trolling identified in RE and SF as trolling behaviour, and essentially classified these in to seven types:

- (1) Engaging by responding sincerely;
- (2) Ignoring the trolling attempt overtly or covertly;
- (3) Exposing the troller to the rest of the group;
- (4) Challenging the troller directly or indirectly;
- (5) Critiquing the effectiveness, success, or 'quality' of the troller;
- (6) Mocking or parodying the trolling attempt; and,
- (7) Reciprocating in kind by trolling the troller.

When we return to the response framework suggested by Harris *et al.* (1986) and later developed by Culpeper *et al.* (2003: 1562–5), most response types could not be accounted for well. In reality, it quickly became clear that, rather than opting for one response or another, users were frequently mixing strategies (e.g., acceptance and defensive, defensive and offensive, or even acceptance, defensive and offensive). It is only fair to note, however, that this same remark can be made about the seven response types given above. Some users reciprocated using mockery, whilst others challenged through critiques, and so forth. That said, these response types do enable us to cast light on the increasing level of face-threatening retaliation that attacked users and (alleged) trollers employed. We can, therefore, emulate that model to an extent, and summarise the outcomes of (perceived) trolling attempts. Trolling can be:

- (1) *Successful*: users do not perceive an intent to troll and are provoked into responding as the troller desires (i.e., they *engage*), risking a high degree of face damage;
- (2) *Frustrated*: users interpret an intent to troll and do not respond (i.e., they *ignore*), risking little or no face damage;
- (3) *Thwarted*: users interpret an intent to troll and counter in a way that reduces the troller's success, or even causes them to lose face (i.e., they *expose, challenge, critique, mock* or *reciprocate*), risking anywhere from none to maximum face damage and possibly further offline consequences; or,
- (4) *Failed*: users both do not perceive an intent to troll and are not provoked. This is included for the sake of completeness, but there is no evidence of this in the data.

There are issues to bear in mind, however. Previous papers (Hardaker, 2010, 2013) have discussed (*i*) that this research is primarily based on examples that Hs claim to perceive as trolling, without being able to account for issues such as mistakes and deception; (*ii*) the value of

quantitatively processing these strategies when faced with combinations of multiple response types; and (iii) the (ever-increasing) semantic scope of 'trolling' as a term, which is as contextually bound and as relative as a term such as 'impoliteness' (Hetcher, 2004; and Opp, 1982; 2001).

A fourth issue that comes to the fore here, is that whilst this paper has covered some response types, Usenet is only one form of CMC, and only two groups have been selected which are, in turn, quite different. RE and SF alone cannot represent all of Usenet – indeed, Usenet newsgroups, even those on the same topic, can contrast markedly in their norms and limits, so we simply cannot assume that all interaction within even one type of CMC will be homogenous, let alone the whole of the Internet. As such, these strategies and responses should only be viewed as an indicative first step towards understanding trolling, and not as an exhaustive list. CMC, by its very nature, is enormous, complex, varied and quickly evolving. Thus, it is more likely than not that analysis of further data will reveal more response types than are presented here, and, indeed, an exhaustive list of response types even with infinitely more data is probably as unattainable as an exhaustive list of any given pragmatic strategy. This does not render these results valueless, however. They provide some insight, however tentative, into long-form, unmoderated, group-based interactions. Further, they stand as useful hypotheses against which to test and explore new data, and can then be extended or discarded on the basis of newer findings. Moreover, this work is an addition, however small, to an extremely under-researched field.

In practical terms, this work also allows for further exploration of the ways in which corpus methods may facilitate pragmatic analyses, despite the apparent mismatch between the two. With regard to forensic linguistics, and particularly forensic pragmatics, this work also draws attention to a range of issues, from determining intent based on what may be deceptive interactions, to deciding upon an instigator when individuals may have manipulated others into overtly attacking first, and to advising individuals of the risks associated with responding in certain ways.

In conclusion, trolling is a far more complex matter than merely launching unprovoked attacks on others; it is open to criticisms of its quality, effectiveness and success, and it can be turned around on the troller so that she is reconstructed as the naive, gullible victim instead of the successful aggressor. Some response types are exactly what a troller would seek and, therefore, only encourage the situation further, whilst others (arguably closer to strategies) appear to be more successful at effectively ending trolling with lesser degrees of risk, whilst others seem to run a small, but serious risk of provoking extreme reactions from trollers who then seek to cause off-line damage.

One thing is clear from this work: far more research is yet needed into linguistic manipulation and deception, and in behaviours such as trolling, cyberbullying, e-hoaxing, and so forth. This paper seeks to take a small step in that direction, but much more still needs to be done.

References

- Andersson, L. and C. Pearson. 1999. 'Tit for tat? The spiraling effect of incivility in the workplace', *Academy of Management Review* 24 (3), pp. 452–71.
- Arundale, R.B. 2008. 'Against (Gricean) intentions at the heart of human interaction', *Intercultural Pragmatics* 5 (2), pp. 229–58.
- Baker, P. 2001. 'Moral panic and alternative identity construction in Usenet', *Journal of Computer-Mediated Communication* 7 (1). Available online at: <http://jcmc.indiana.edu/vol7/issue1/baker.html>
- BBC. 2010. 'Jade Goody website "troll" from Manchester jailed.' BBC News, 29 October 2010. Accessed 12 November 2010 at: <http://www.bbc.co.uk/news/uk-england-manchester-11650593>
- Bernstein, M.S., A. Monroy-Hernández, D. Harry, P. André, K. Panovich and G. Vargas. 2011. '4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community', *Association for the Advancement of Artificial Intelligence*, pp. 1–8.
- Binns, A. 2011. 'Don't feed the trolls: managing troublemakers in magazines' online communities.' *Mapping the Magazine* 3 (Cardiff University), pp.
- Camber, R. and S. Neville. 2011. 'Sick internet 'troll' who posted vile messages and videos taunting the death of teenagers is jailed for 18 WEEKS.' *Daily Mail*, 14 September 2011. Accessed 2 December 2011 at: <http://www.dailymail.co.uk/news/article-2036935/Natasha-MacBryde-death-Facebook-internet-troll-Sean-Duffy-jailed.html>
- Carney, T. 2014. 'Being (im)polite: a forensic linguistic approach to interpreting a hate speech case', *Language Matters* 45 (3), pp. 325–41.
- Chakraborti, N. 2010. *Hate Crime: Concepts, Policy, Future Directions*. London: Willan.
- Culpeper, J., D. Bousfield and A. Wichmann. 2003. 'Impoliteness revisited: with special reference to dynamic and prosodic aspects', *Journal of Pragmatics* 35 (10–11), pp. 1545–79.
- Culpeper, J. and C. Hardaker. Forthcoming. 'Pragmatics and corpus linguistics' in P. Baker and J. Egbert (eds) *Triangulating Methodological Approaches in Corpus Linguistic Research*. London: Routledge.

- Donath, J.S. 1999. 'Identity and deception in the virtual community' in M.A. Smith and P. Kollock (eds) *Communities in Cyberspace*, pp. 29–59. London: Routledge.
- Douglas, K.M. and C. McGarty. 2001. 'Identifiability and self-presentation: computer-mediated communication and intergroup interaction', *British Journal of Social Psychology* 40 (3), pp. 399–416.
- Doyle, J. 1989. 'Re: <hick!>.' alt.callahans 14th December, pp. <https://groups.google.com/d/msg/alt.callahans/SphfCkUsdtY/FggkP4rvoQEJ>.
- Facebook. 2010. 'Statement of rights and responsibilities'. Accessed 13 May 2010 at: <http://www.facebook.com/terms.php>
- Felson, R.B. 1982. 'Impression management and the escalation of aggression and violence', *Social Psychology Quarterly* 45 (4), pp. 245–54.
- de Fina, A. 2006. 'Group identity, narrative, and self-representations' in A. de Fina, D. Schiffrin and M. Bamberg (eds) *Discourse and Identity*, pp. XX. Cambridge: Cambridge University Press.
- Fraser, B. 1998. 'Threatening revisited', *Forensic Linguistics* 5 (2), pp. 159–73.
- Gibbs, R.W. 2001. 'Intentions as emergent products of social interactions' in B. Malle, L. Moses and D. Baldwin (eds) *Intentions and Intentionality*, pp. 105–122. Cambridge, Massachusetts: MIT Press.
- Golder, S.A. and J.S. Donath. 2004. 'Social roles in electronic communities' in Association of Internet Researchers (AoIR) Conference: Internet Research 5.0, pp. 1–25. 19–22 September. Brighton, England.
- Grimshaw, A.D. (ed.). 1990. *Conflict Talk: Sociolinguistic Investigations of Arguments in Conversations*. Cambridge: Cambridge University Press.
- Hardaker, C. 2010. 'Trolling in asynchronous computer-mediated communication: from user discussions to academic definitions', *Journal of Politeness Research: Language, Behaviour, Culture* 6 (2), pp. 215–42. doi: 10.1515/jplr.2010.011.
- Hardaker, C. 2013. "'Uh.....not to be nitpicky,,,,,but...the past tense of drag is dragged, not drug": an overview of trolling strategies', *Journal of Language Aggression and Conflict* 1 (1), pp. 57–85.
- Harris, L., K. Gergen and J. Lannaman. 1986. 'Aggression rituals', *Communication Monographs* 53 (Sept.), pp. 252–65.
- Haugh, M. 2008. 'Intention in pragmatics', *Intercultural Pragmatics* 5 (2), pp. 99–110.

- Haugh, M. 2010. 'Jocular mockery, (dis)affiliation and face', *Journal of Pragmatics* 42 (8), pp. 2106–19.
- Herring, S.C. 2003. 'Computer-mediated discourse' in D. Schiffrin, D. Tannen and H.E. Hamilton (eds) *The Handbook of Discourse Analysis*, pp. . Oxford: Blackwell.
- Herring, S.C., K. Job-Sluder, R. Scheckler and S. Barab. 2002. 'Searching for Safety Online: Managing 'Trolling' in a Feminist Forum', *The Information Society* 18, pp. 371–84.
- Hetcher, S.A. 2004. *Norms in a Wired World*. Cambridge: Cambridge University Press.
- Jucker, A.H. 2013. 'Corpus pragmatics' in J.-O. Östman and J. Verschueren (eds) *Handbook of Pragmatics*, pp. 1–17. Amsterdam: Benjamins.
- Jucker, A.H., D. Schreier and M. Hundt (eds). 2009. 'Corpora: pragmatics and discourse', *Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME 29)*. (Language and Computers: Studies in Practical Linguistics 68). Ascona, Switzerland, 14–18. Amsterdam: Rodopi.
- Labov, W. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Oxford: Blackwell.
- Lampe, C. and P. Resnick. 2004. 'Slash(dot) and burn: distributed moderation in a large online conversation space' in proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves pp. 543–50. New York, USA: ACM.
- Leech, G.N. 1983. *Principles of Pragmatics*. London: Longman.
- Lein, L. and D. Brenneis. 1978. 'Children's disputes in three speech communities', *Language in Society* 7 (3), pp. 299–323.
- Maddox, T. 1989. 'Re: Cyberspace Conference.' alt.cyberpunk 22nd October, pp. <https://groups.google.com/d/msg/alt.cyberpunk/976Vj9FPX3Q/3Ytxg-JeCdMJ>.
- Mauney, J. 1982. 'second verse, same as the first.' net.nlang 5 July, pp. https://groups.google.com/d/msg/net.nlang/YNX8PIANL_g/Uveya3fB1ZkJ.
- McLaughlin, M.L., K.K. Osborne and C.B. Smith. 1995. 'Standards of conduct on Usenet' in S. Jones (ed.) *CyberSociety: Computer-mediated Communication and Community*, pp. 90–111. Thousand Oaks, California; London: Sage Publications.
- Miller, M. 1990. 'FOADTAD.' alt.flame 8 February, pp. <https://groups.google.com/forum/?fromgroups=#!msg/alt.flame/RMlMz6ft4r8/SwPcwiHXE4AJ>.

- Mills, S. 2003. *Gender and Politeness*. Cambridge: Cambridge University Press.
- Mooney, A. 2004. 'Co-operation, violations and making sense', *Journal of Pragmatics* 36 (5), pp. 1601–23.
- Morris, S. 2011. 'Internet troll jailed after mocking deaths of teenagers.' *Guardian* 13 September 2011. Accessed 15 September 2011 at: <http://www.guardian.co.uk/uk/2011/sep/13/internet-troll-jailed-mocking-teenagers>
- O'Keefe, B.J. 1989. 'Communication theory and practical knowledge' in B. Dervin (ed.) *Rethinking Communication*, pp. 197–215. Newbury Park: Sage.
- Opp, K.-D. 1982. 'The evolutionary emergence of norms', *British Journal of Social Psychology* 21, pp. 139–49.
- Opp, K.-D. 2001. 'How do social norms emerge? An outline of a theory', *Mind and Society* 2 (1), pp. 101–28.
- Plato. 2007. *The Republic*. (Third edition.) London: Penguin Classic.
- Romero-Trillo, J. (ed.). 2008. *Pragmatics and Corpus Linguistics*. Berlin: Mouton de Gruyter.
- Scott, M. 2009. *WordSmith Tools*. Liverpool: Lexical Analysis Software.
- Shachaf, P. and N. Hara. 2010. 'Beyond vandalism: Wikipedia trolls.' *Journal of Information Science* 36 (3), pp. 357–70.
- Shin, J. 2008. 'Morality and internet behavior: a study of the Internet troll and its relation with morality on the Internet' in K. McFerrin, R. Weber, R. Carlsen and D. A. Willis (eds) *proceedings of Society for Information Technology and Teacher Education International Conference 2008*, pp. 2834–40. Chesapeake, VA: AACE.
- Sia, C.-L., B.C.Y. Tan and K.-K. Wei. 2002. 'Group polarization and computer-mediated communication: effects of communication cues, social presence, and anonymity', *Information Systems Research* 13 (1), pp. 70–90.
- Siegel, J., V.J. Dubrovsky, S. Kiesler and T.W. McGuire. 1986. 'Group processes in computer-mediated communication', *Organizational Behaviour and Human Decision Processes* 37 (2), pp. 157–87.
- Taavitsainen, I., A.H. Jucker and J. Tuominen (eds). 2014. *Diachronic Corpus Pragmatics*. Amsterdam: John Benjamins.
- Tepper, M. 1997. 'Usenet communities and the cultural politics of information' in D. Porter (ed.) *Internet Culture*, pp. 39–54. New York: Routledge.
- Utz, S. 2005. 'Types of deception and underlying motivation: what people think', *Social Science Computer Review* 23 (1), pp. 49–56.

- Vinagre, M. 2008. 'Politeness strategies in collaborative e-mail exchanges',
Computers and Education 50 (3), pp. 1022–36.
- Watts, R.J. 2003. Politeness. Cambridge: Cambridge University Press.

Table 1: Comparative information concerning RE and SF

Token	RE	SF
Group creation date	2 Dec. 87	4 June 98
Corpus collection date	31/05/11	31/05/11
Group subscriber-count at corpus collection date	1,106	211
Group post-count at corpus collection date	795,349 (avg 92.767 per day)	288,163 (avg 60.472 per day)
Corpus date-range	1 July 05 to 30 June 10	10 Mar. 05 to 30 June 10
Corpus post-count	170,634	57,734
Corpus word-count ⁶	62,884,032	16,596,565

⁶ Wordcount is a very weak guide, since Usenet posts frequently (re)quote, and discounting (re)quoted material to acquire a real wordcount, whether automatically or manually, is very cost-benefit prohibitive.