UNIVERSITY OF HELSINKI

https://helda.helsinki.fi

Negotiating offensive humour online : 'Only joking' claims on Reddit

Biri, Ylva; Tanskanen, Sanna-Kaisa

2025-09-26
John Benjamins
http://hdl.handle.net/10138/626184
Biri, Y & Tanskanen, S-K 2025, 'Negotiating offensive humour online : 'Only joking' claims on Reddit', Pragmatics & Cognition, vol. 32, no. 1, pp. 151-177. https://doi.org/10.1075/pc.24023.bir

Downloaded from Helda, University of Helsinki institutional repository. https://helda.helsinki.fi
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.

This is the Author's Accepted Manuscript of the following article:

Biri, Ylva, and Sanna-Kaisa Tanskanen. 2025. 'Negotiating Offensive Humour Online: "Only Joking" Claims on Reddit'. *Pragmatics & Cognition* 32 (1): 151–77. https://doi.org/10.1075/pc.24023.bir.

The article is part of the special Issue "Negotiating offensive humour in interaction" edited by C-H. Elder, E. Kapogianni & I. Baxter-Webb.

Negotiating offensive humour online: 'only joking' claims on Reddit

Ylva Biri, University of Helsinki

Sanna-Kaisa Tanskanen, University of Helsinki

This paper investigates the negotiation of offensive humour through 'only joking' claims made by participants on the social media platform Reddit. Metapragmatic claims of jocular intent can be added to a post containing offensive humour pre-emptively or after a reaction by another participant. In both cases, the participant making the claim tries to show that problematic humour in the post is not meant to give offence, and that they should not be held accountable for such an intention. Using corpus methods and close analysis, we analyse indicators of jocular intent (e.g. "just kidding", "take a joke") to find out how offensive or problematic humour is presented, perceived and negotiated. The findings show how claims of jocular intent can be accepted or rejected after a negotiation of the offensiveness based on subreddit community norms. Offensive humour is negotiated verbally and through the upvote/downvote affordance of Reddit, indicating the ascription of accountability.

Keywords: metapragmatics, intentionality, accountability, jocular intent, offence, Reddit

1 Introduction

Offence can be taken even when it was not intended, which may result in a negotiation of who is at fault (Dynel 2021; Elder 2021a). This paper explores metapragmatic claims of jocular intent on an online platform, such as "only joking" or "just kidding", through which a participant can defend their (or another participant's) post that may cause or has caused offence. By successfully defending their intent as humorous, the offending party can reject their accountability of causing offense and instead sanction the offended party for having been offended (Haugh 2016; Chang & Haugh 2021; Skalicky, Berger & Bell 2015). However, a defence based on there having been no intent to offend does not guarantee clearance: people can feel hurt or annoyed even knowing that no offence was intended by the speaker (Culpeper 2011; Haugh 2015).

Our goal is to investigate how claims of jocular intent are presented, perceived, contested and negotiated by users in online interaction. Claims of jocular intent are metapragmatic expressions that serve as a starting point for investigating the negotiation of intent, which other interactants can contest or defend. The negotiations display reflexive or metapragmatic awareness of the participants, that is, the interactants' emic understanding of language use and interaction (Haugh 2018; Sinkeviciute 2019; Verschueren 2000). As such, our analysis, too, defines 'offensive jokes'

using the interactants' perceptions and knowledge of the local culture as a starting point. We acknowledge that what might be sanctioned as offence in one context may well be welcomed as a joke or playful teasing in another. However, a joke that participants feel the need to defend as "just a joke" likely has the potential to be misconstrued at least by some of the target audience.

With online interaction, we bring a new context to the investigation of jocular claims, which have to date mainly been studied in various types of spoken interaction. Our research context is Reddit, a discussion forum and social news aggregation website. The platform hosts thousands of subforums, or *subreddits*, which are often conceptualised as separate online communities, whether based on weak or strong ties; each subreddit has its own discussion topics, rules, expectations, and rituals (Robards 2018). These subreddits function as local contexts of humour, offence and negotiating responsibility for giving or taking offence. Our dataset of authentic online comment sequences containing claims of jocular intent is thus analysed for how instances of offensive humour are negotiated by Redditors drawing on the community norms of Reddit and the relevant subreddit. We argue that community norms are central for what kind of humour participants consider permissible and for how accountability of offense is negotiated, whether or not this discussion leads to a consensus. In addition to verbal means, participants can use the non-verbal upvote/downvote affordance of Reddit in the negotiation.

In the following, we first present the background, data and methods of the study. In the analysis, we show (RQ1) how Reddit users clarify their jocular intent either pre-emptively or reparatively after a reproach, and (RQ2) how they defend the jocular intent of another user. We also pay close attention to (RQ3) how Redditors use upvotes and downvotes to signal their (dis)approval of jocular claims. The analysis concludes with examples of the use of another Reddit affordance, namely editability.

2 Jocular claims, intentions and accountability

In this paper, we investigate claims of jocular intent made in connection with posts on the social media platform Reddit. But why would users make such claims on Reddit, or anywhere else, for that matter? That such claims can be found in interaction is an indication of a (potential) problem or trouble in interaction. The need to clarify intent arises when a user considers that there is potential for a problem or misunderstanding, which in our case means that the jocular intent must be clarified in a pre-emptive claim. Clarifying jocular intent may also become necessary after it has emerged that other users (recipients) have either not recognised the jocular intent or they reject it (cf. Elder & Haugh 2023 on unwanted inferences).

Starting with Grice (1957), researchers have repeatedly noted that an utterance can produce numerous interpretations (Elder & Haugh 2023; Hansen & Terkourafi 2023). This is because recipients draw inferences based on what the speaker (writer) has said, but going beyond what was said to what may have been intended (implicatures). Hence, the recipients' interpretation may not match the original intention of the speaker. In addition, Haugh (2007) questions how consciously intentions are formed in the speaker's mind at the time of speaking and whether they can indeed be clear and predetermined. Discarding the premise of fixed speaker intentions opens the possibility of interpretations that may not match the speaker's original, potentially vague intent (Elder 2021b).

Negotiating intent is thus interactional and dynamic. Haugh (2007) maintains that speaker and recipients can together create implicatures, regardless of the speaker's initial intentions; in other

words, speaker and recipients can negotiate meanings in interaction (Elder & Haugh 2018). Furthermore, instances where the speaker's intention and the recipient's interpretations do not align but still lead to successful communication, broaden the traditional notion of effective communication (Elder 2021b). This does not mean, however, that misunderstandings would not arise or miscommunication happen (see e.g. Dynel 2017); in some cases, the mismatch between the speaker's intentions and recipients' interpretations can be too wide.

To address the challenge of understanding speaker's intentions and recipients' interpretations, Haugh (2008) suggests examining how intentions are negotiated after the fact. When language users attempt to clarify their intended message or challenge an attributed meaning *post facto*, they engage in metapragmatic negotiation. How jocular claims are used in such negotiation has previously been investigated in teasing sequences in conversational interaction in several languages. The studies have uncovered several functions or "forms of moral work" (Haugh 2016: 133) for jocular claims. These include blocking the taking of offence through pre-emptive disclaimers, repairing failed humour, sanctioning a recipient for taking the joke too seriously, or signalling a return to serious talk (Haddington 2011; Haugh 2016; Chang & Haugh 2021; Skalicky et al. 2015).

Similarly, recipients can argue for their interpretations of the intent. Through this metapragmatic negotiation, intentions are often intertwined with questions of accountability, as language users are held accountable for what they have expressed (Elder 2021a; Garfinkel 1984; Haugh 2008). Hansen and Terkourafi (2023) highlight the importance of speakers being held accountable for the meanings ascribed to their utterances. Accountability extends to recipients' interpretations, meaning that speakers might have to answer for inferences they never meant to provoke (see also Elder 2021a; Dynel 2021; Mitchell & Haugh 2015).

We understand metapragmatics as the display of reflexive awareness by language users of their use of language (see e.g. Caffi 1998; Haugh 2018; Verschueren 2000). More specifically, we focus on metapragmatics in use, which means analysing metapragmatic acts which are "put to communicative practice and by which interactants intervene in ongoing discourse" (Hübler & Bublitz 2007). These acts are a visible sign of the "management of discourse" (Caffi 1998), whereby interactants monitor and reflect on their own and other people's communicative behaviour. They are one way of signalling the understanding that interactants have of how language is used for interaction (Blommaert & Rampton 2011; Haugh 2018; Verschueren 2000). Metapragmatic acts offer the participants' own, emic perspective on the interaction (Dynel & Poppi 2019; Liu & Liu 2021; Sinkeviciute 2019). The acts can be used to fulfil various functions, but their reflexive basis means they are inherently evaluative. There has been active research on metapragmatic negotiations in online interaction in recent years, and the features of language use studied range from appropriateness and identity to rudeness and conflict, among others (see Bou-Franch 2021; Dynel & Poppi 2019; Kleinke & Bös 2015; Tanskanen 2007; 2018; 2021; 2024).

In the context of the present study, by making a jocular claim, users are informing others that the intent of their own post or a post by another user is humorous; this is the meaning they should be seen to be communicating. The intent and the claim can be disputed by other users, and other interpretations offered (including offence), for which the poster is then held accountable. Example (1) from our Reddit data illustrates how jocular intent is metapragmatically negotiated. The subreddit in question is dedicated to discussion on aquariums, and the original post is a photo from a pet shop, which shows a new line of furniture and products for pets called Nate + Jeremiah, advertised with a photo of two smiling men.

(1) A: Nate and Jeremiah? Sounds like a Mormon law firm. I bet their wives are enjoying the cash flow though.
B: I'm no expert but I'm pretty sure nate and jeremiah are married to each other
A: I thought about putting the /s tag. But I thought people would get the joke. I was wrong.
B: It didn't really seem like a joke, but maybe I'm just embittered as a gay ex-mormon. Have a nice day/night

User A's comment on the original post sparks a reply from User B, who questions the accuracy of A's post. Following B's reaction, to repair the failed joke, A returns to clarify the jocular intent of their post, explaining that they thought the intent would have been clear without a metapragmatic /s tag (for sarcastic humour). In other words, A has actively decided against using a pre-emptive disclaimer, trusting that the humorous intent would be evident (cf. Dynel 2023: 242). While A's reply could be regarded as sanctioning B for taking the joke too seriously, by adding "I was wrong", A acknowledges the interpretation made by B and indicates willingness to be held accountable for a joke that has misfired. Although it does not signal complete acceptance of the jocular claim, B's final reply displays similar acknowledgement of A's original intent. In their negotiation, both A and B display "reflexive accountability" (Elder 2021b: 65); they acknowledge and react to each other's clarifications, and in the end arrive at an understanding that jocular intent was present, although it may not have been clear. All in all, despite the divergence between A's intent and B's interpretation, the two Redditors manage to interact successfully.

3 Offensive humour in online communities

Interactional intent is understood in context and the frame or discourse types of an interaction event help interlocutors understand each other's probable intent; local practices guide the expectations of conventional meanings and moral norms of appropriateness (Hansen & Terkourafi 2023; Tagg, Seargeant & Brown 2017). An online community with shared genre norms and specialised lexis often constitutes a 'discourse community' (Swales 1990: 24–27) or even a 'community of practice' with jointly identified goals and activities (Lave & Wenger 1991). In a heterogeneous online forum such as Reddit, however, it is unlikely that all interactants have a shared understanding or interpretation of an utterance. Mismatched expectations and norms may result in misreading of the situation and in unintentional offence. This is indeed what causes the need for clarifying intent. Yet correct interpretation of intent is possible if participants make correct assumptions about the other interactants and any explicit or implicit community norms. Therefore, both (alleged) attempts at humour and offence can and should be interpreted and evaluated based on community norms and emic interpretations.

Online interaction is infamous for offensive humour such as provocative content and humorous memes, which blur the line between offensiveness and humour. Among research on distinguishing offensive and jocular content, Dynel (2021a, 2021b) identifies intent as the key for separating genuine insults from jocular insults; the former intend to offend whereas the latter intend to amuse. Offensive intent is present online, for example, in the deliberately provocative 'trolling', sometimes amounting to cyberbullying or hate speech; to its practitioners, trolling may be "just some game in the online community" (Erjavec & Kovačič 2012: 912; also Arendholz 2013: 121). In this reframing of offensive content as a game, causing offence equals to winning the game whereas taking offense means losing or misunderstanding the game. In contrast, the intent to amuse and joking as an activity is successful when all parties involved recognise the jocular frame,

even when the joke relies on subversive or provocative humour. For example, jocular activities such as banter, teasing or roasting require that interlocutors recognise that what appears aggressive on the surface is in fact benevolent (Dynel 2008; Dynel & Poppi 2019). In example (2) from r/AskReddit, User A acts as the 'Original Poster' (OP) by initiating a discussion thread about geniuses. User B suggests that Redditors are less intelligent than they think, and User A responds with a phrase that could be read as a threat ("you got a fight coming").

(2) A (OP): Where are all the geniuses at?
B: On Reddit, mouthing off half-answers to questions that are actually above their heads. (+2)
A: Oh boy have you got a fight coming. Jk I'm dumb. (+2)

User A and B can be seen as participating in self-deprecating humour: User A's jocular intent is clarified with "Jk", short for "just kidding", combined with the self-deprecating "I'm dumb". Both A and B's comments have been upvoted, further suggesting a recognition of the well-intentioned jocular genre. Joking about the illusory superiority of Reddit users on Reddit is thus recognised as teasing that pokes fun at ingroup members, including A and B themselves. Community norms regulating teasing activities can also be more specific and explicit, as in the case of Reddit communities that are devoted to humorous roasting but do not tolerate genuine insults (Allison, Bussey & Sweller 2019; Dynel & Poppi 2019). The community's shared understanding of the humorous practices ensures that an intent that on the surface could be construed as offensive is understood as humorous.

Subversive and taboo humour is often used to boost affinity and solidarity within a likeminded online community. In instances of aggressive or norm disruptive humour, the joke-teller's intent may be to ridicule or insult outgroups while amusing ingroups, thereby strengthening ingroup solidarity. For example, jokes and memes about societal and political issues are often based around negative stereotypes about ideological opponents while strengthening the community's or ingroup's solidarity and shared values and evaluations (Altahmazi 2024; Andersson 2021; Dynel 2020; Zappavigna 2022). Such degrading humour and parody is also used to mark affiliation in social media criticism or protests of more mundane topics such as customer experience (Orthaber 2019) or moderation in a social media group (Mullan 2022). Also, humour based around taboo topics such as death, war and serious diseases may boost solidarity. Dark humour during the COVID-19 pandemic was circulated in online memes to help individuals cope with the quarantine (Aslan 2022). Similarly, the humorous memes tweeted by the government of Ukraine after the invasion by Russia act both as "digital warfare weapons" and indicators of resilience (Dynel 2024: 254). As a final example, in a peer-support forum for cancer survivors and fighters, humour about cancer is empowering for the participants' shared identities as survivors (Demjén 2016). Even though joking about such topics is generally not socially acceptable, it can be locally acceptable. In sum, both aggressive and taboo humour for strengthening solidarity are aimed at an ingroup that understands and appreciates the purpose of the joke, whereas the parties that might take offense – or who are not permitted to joke about the topic – are outgroup members.

The local expectations and practices influence whether and what kind of provocative content is permitted and when. Both the sender and recipient of a joke online must make use of context specific norms of what is or could be seen as offensive and/or amusing. Unsuccessful humour and miscommunication in the form of offence often stems from differing expectations or appropriateness in intradiverse online contexts (Tagg, Seargeant & Brown 2017: 87). As such, we

account for the online community's own norms and interpretation in our analysis of Reddit users' metapragmatic awareness and negotiation of intent.

4 Data and methods

4.1 Reddit as a research context of offensive humour

Reddit is a popular pseudonymous social media and news aggregation platform. It is divided into more than 100,000 active 'subreddits' (and hundreds of thousands more inactive subreddits), each defined by a conversation topic or purpose. This makes Reddit a fruitful data source for charting the negotiation of offensive humour in different contexts and in communities of varying degrees of homogeneity. Many subreddits are devoted to humorous content in general, including r/funny and r/memes, which at 66 million and 35 million subscribers[1] respectively are among the most popular subreddits. Yet others are reserved for general discussion on specific topics (e.g., r/worldnews (44M subscribers) or r/science (34M subscribers)) or specific social groups (e.g., r/teenagers (3.1M subscribers), r/electricians (480K subscribers)); in these subreddits, humour might be less expected or less welcome depending on the nature of the subreddit. Even so, given that humour is an inherent part of social interaction and affiliation, it can be found in discussions on practically all topics and irrespective of subreddit size.

Reddit makes for an interesting research context for offensive humour also because of its decentralised moderation. Reddit has a general content policy as well as a Reddiquette, a regularly updated document "written by redditors themselves", which forbids, for instance, spamming, harassment, trolling, ad hominem attacks and "[being] (intentionally) rude" (Reddiquette, 2025); note how these platform rules list not only problematic content but also intent to offend as forbidden. In addition to the Reddiquette, individual subreddits may set their own rules according to the community's own goals and interests, for example, to keep discussions on topic. The interpretation of (in)appropriateness is thus defined also at a local community level and communities may have conflicting norms of what is humorous or what is offensive. For example, while Reddit has been criticised for heteronormative misogyny (e.g., Massanari 2017), it also homes communities devoted for feminist and LGBTQIA+ topics (e.g. Robards 2018). Sharing and browsing content on a platform with such a heterogeneous audience leads to potential context collapse (Marwick & boyd 2011). Although misunderstandings of communicative intent may naturally occur also between closely aligned individuals, we expect the varying community norms to result in misunderstandings, offence, and thereby the need for clarification and negotiation of intent.

Although Reddit is based around verbal, written asynchronous interaction, Reddit posts can be upvoted or downvoted by any user: each Reddit post starts off with the default score of +1, which is then adjusted based on the votes it receives.[2] Upvotes and downvotes are 'paralinguistic digital affordances' (Hayes, Carr & Wohn 2016), non-verbal acknowledgements with upvotes generally signalling agreement or approval and downvotes disapproval. Up- and downvotes also determine

[1] Subreddit subscribers, or 'members' in Reddit's own term, refers to users who have subscribed to content posted in a subreddit. However, because subscription or membership is not required for browsing or posting in a subreddit, the membership numbers do not necessarily give an accurate picture of a subreddit's level of activity.

[2] The score is the sum of upvotes and downvotes, but the exact number of each type of vote is unavailable in the metadata. Additionally, there is no data on which users vote which way. Thus, a high score in either direction can be read as indicating community opinion, but a low score is only tentative.

the visibility of the content, with highly upvoted content listed higher up. This means that upvoting is a way to promote answers deemed particularly helpful, insightful or funny. Of course, upvotes may create a feedback loop where increased visibility means more attention and more upvotes (e.g. Glenski et al. 2018). This means that a comment with a small number of upvotes is not necessarily 'poor', it may simply not have gained enough traction to be read by a large audience. Nevertheless, the up- and downvotes in combination with close reading can be used to gauge audience reaction and perlocutionary success.

4.2 Data collection and analysis

Our analysis uses corpus pragmatics to automatise the identification of 'only joking' claims in a large dataset of Reddit comments. By searching for forms associated with an illocutionary function, we can scale the search on a large dataset, which is necessary especially when the dataset is substantial, and the function of interest is quite infrequent (Kohnen 2007; O'Keeffe 2018). Our operationalisation focuses on metapragmatic acts that are made explicit through a set of phrases. It excludes jokes where users did not feel the need to clarify their intent but also situations where jocular intent might be clarified through other means.

Our list of phrases associated with clarifying jocular intent was based on an initial search of typical phrases associated with the lemmas *joke*, *joking*, and *kidding*. Similar to Skalicky, Berger and Bell's (2015: 22) corpus-assisted analysis, variants of "just kidding" (including the acronym "jk") were the most frequently used, followed by variants of "just joking". In addition, the following phrases were searched for: "only joking", "only kidding", "BE a joke", "just a joke", "take a joke", "GET the/a joke", "BE joking/kidding", "kidding around", "joking around", as well as "joke lol", "joking lol", "kidding lol". The phrases with the internet slang term *lol* were included because *lol* was found to regularly co-occur with *joking* and *kidding*, indicating its use as an illocutionary marker to further mark a non-serious or affiliative meaning (Schneebeli 2020). We also searched for the metapragmatic tag "/jk" or "/j", inserted at the end of a comment to indicate that the preceding text is a joke. As we focus on jokes, metapragmatic labels and tags referring to other non-serious intent (e.g. "/s" for sarcasm) were excluded.

We accessed Reddit comments from October 2022 through the Pushshift Reddit Dataset (now defunct) (Baumgartner et al. 2020) and retrieved instances of the phrases associated with jocular intent. From the first 2 million comments in the total dataset, we found a total of 1781 instances containing any of the search terms: these were retrieved and saved as concordance lines into a spreadsheet, where they were annotated for their function. We have deindividualized the examples quoted in this paper. However, the data was publicly available at the time of analysis.

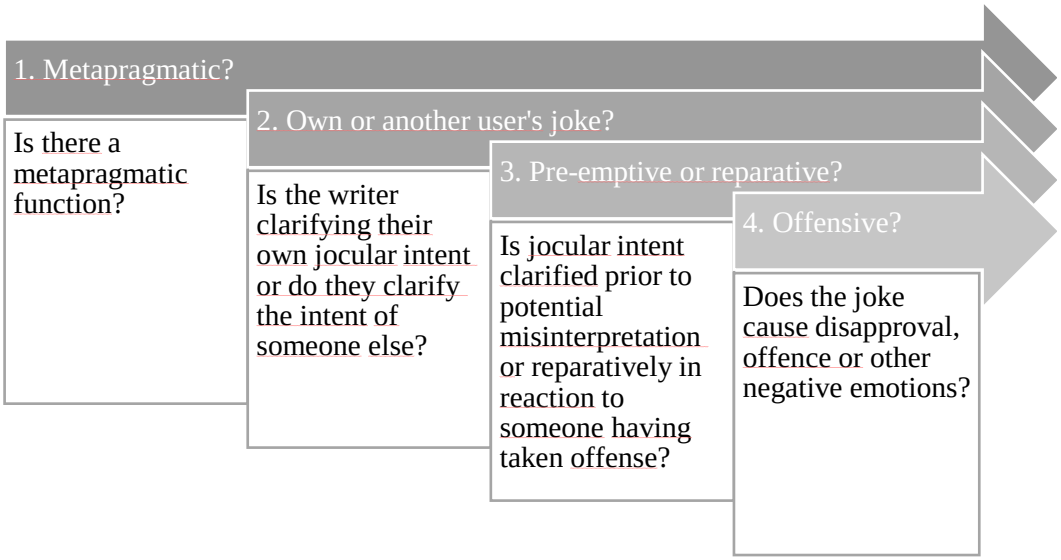

Figure 1: Schema for annotating metapragmatic claims of jocular intent

Figure 1 illustrates the annotation schema of determining whether and how a concordance line is associated with the negotiation of jocular intent and accountability. We first analysed (1) whether the instance represents a metapragmatic function of the search phrase. Instances that were not metapragmatic were discarded from further analysis.[3] While the metapragmatic use of the phrase could often be assessed in the limited keyword-in-context concordance lines, for the other three steps, we consulted the original context on Reddit providing that it was still accessible online. Instances identified as metapragmatic claims were further analysed in terms of (2) whether the writer is clarifying their own or someone else's humorous intent. In step (3), we analysed whether the jocular intent is clarified pre-emptively or reparatively, that is, whether the claim occurs in the same message as the (potentially problematic) jocular content, before offence could be taken, or whether it is a reaction to someone having expressed their taking offence.

In line with our view of intent and metapragmatic acts being situated in a community context, we define 'causing offence' and 'offensiveness' based on the emic view or observed reactions (step 4 in Figure 1). Crucially, offensiveness is not defined or classified based on the speaker's (writer's) intent nor the topic or butt of the joke but rather based on whether it causes offence among the hearers (readers) (see Hansen & Terkourafi 2023). If a writer's attempt at humour has sparked offence, annoyance, or other negative emotions or moral judgements among other interlocutors, it is classified as offensive. Through a pre-emptive clarification of a comment being "just a joke" the writer indicates awareness of a (potential) problem, but from an emic perspective it does not need to be marked as offensive unless it has actually caused some grievance. Offence may be evident in verbal criticism or evaluation but also non-verbally in the form of downvotes signaling a negative evaluation of the quality or appropriateness of the joke. It is this kind of response that a metapragmatic act can seek to repair. As will be seen in the analysis, the audience is not necessarily unanimous about the offensiveness of a joke, which may make a binary classification problematic. However, usually a general consensus emerges from the users' interpretation and evaluation of a joke and its offensiveness. By operationalising offensiveness as a quality

[3] These included, for example, the use of "jk" as initials and phrases in contexts such as "my life is a joke" or "my wife was joking about it" (referring to narrated past events).

determined by the online community, we as researchers can analyse the negotiation and interaction of users without imposing our own impressions on inside jokes or jocular practices in a context unfamiliar to us.

The authors of this paper annotated 1031 concordance lines and a research assistant instructed in the task annotated the remaining 750 lines. Any unclear instances in both sets were discussed jointly between the authors. The total number of concordance lines was 1781 from across 999 subreddits, of which metapragmatic function was attested in 873 instances from across 579 subreddits (see Table 1). Around ¾ of metapragmatic claims were related to the writer's own joke and intent, with the remaining ¼ clarifying the intent of another user's post (Table 1). Defence of another user's joke is practically inherently a reparative act. However, metapragmatic acts relating to own posts were fairly evenly divided across the pre-emptive and reparative functions; this is in contrast to the data used by Skalicky, Berger and Bell (2015: 22), where pre-emptive use was twice as frequent as repairing functions.

Table 1: Frequency of metapragmatic claims of jocular intent

	Total	% of total
Total number of concordance lines	**1781**	**100.0%**
Metapragmatic, own intent	**681**	**38.2%**
- pre-emptive, non-offensive	314	17.6%
- pre-emptive, offensive	23	1.3%
- reparative, non-offensive	252	14.1%
- reparative, offensive	92	5.2%
Metapragmatic, other's intent	**192**	**10.8%**
- non-offensive	138	7.7%
- offensive	54	3.0%
Non-metapragmatic	**754**	**42.3%**
Unavailable or NA	**154**	**8.6%**
Metapragmatic total	**873**	**100.0%**
- Metapragmatic non-offensive	704	80.6%
- Metapragmatic offensive	169	19.4%

In our analysis, most concordance lines were annotated as non-offensive, that is, as successful jokes at least in the eyes of the majority of the audience. Offence has been taken in 169 instances (19.4% of all metapragmatic functions, see Table 1). The next section will further illustrate how the negotiation of offence can play out and how users seek to clarify a potential misunderstanding of intent – either successfully or unsuccessfully.

5 Analysis

5.1 Defending own jocular intent

5.1.1 Pre-emptive clarifications

Pre-emptive clarification of jocular intent (i.e. a disclaimer) suggests that the user recognises their joke as potentially problematic prior to anyone having expressed offence or having had the chance to do so. Although the reader may nevertheless take offence at a joke in poor taste, clarification of intent in advance can prevent misunderstanding and mitigate communicative failure. The success of a joke and the successful clarification of intent are evident in positive reactions to the joke or in subsequent turns picking up on the joking frame. In example (3), User A is discussing the actions of a shoplifter suspect in *On Patrol: Live*, a reality television series about law enforcement.

(3) A: So does she just like wander around the store and put stuff in her bag? Like … someone's going to see you. You gotta grab and dash. (+ 4)
B: Speaking from experience??? Just kidding... (+1)
A: Haha! Working in retail scars you for life! 😂 (+3)
B: Oh, I'm sure it does (+1)

User B suggests User A, too, might be a shoplifter, but clarifies that they were kidding. User A acknowledges and approves of this intent through a reply containing a laughter particle and the

face with tears of joy emoji, though clarifying that their experience with shoplifters is from a store employee perspective.

Pre-emptive clarifications can also serve as transitions from non-serious to serious (Schegloff 2001; see also Chang & Haugh 2021). The clarification may function as a frame shift from jocular to serious as in (4), where User A as the thread initiator, i.e. the 'OP', has shared a photo of cooked chicken, to which User B reacts with mock-disgust.

(4) A (OP): [photo of cooked chicken with title as caption] Boneless skinless thighs are my favourite
B: Skinless?! Blasphemy! Naw, I'm kidding. They look great! How did you do them? (+28)

User B's shift to a compliment and a genuine question is acknowledged by User A, who proceeds to give a description of their recipe. A's response combined with the high number of upvotes for B can be seen as a successful joke and mitigation of potential misunderstanding. Similar to the teasing in example (2) above, examples (3) and (4) illustrate two interlocutors (and observers) who acknowledge and appreciate the humorous intent, with no indication of offence being taken.

Pre-emptive clarification of jocular intent does not guarantee success. Although it can be seen as a defence against interaction failure such as disapproval or offence (Bell 2015: 112), other users may still deem the joke tasteless or offensive. In example (5), User A requests advice in preparation for their first ever high school party.

(5) A (OP): I'm going to my first high school party tomorrow night. Is there any thing I should know or is there any advice for me? Thanks.
B: Get as drunk and high as you can. I'm only joking… (0)

While User B's joke has received no comments, i.e. no verbal criticism, it has a score of 0. As the default score is +1, this means that at least one user has made use the platform's paralinguistic affordance to indicate they deem the comment unfunny and/or inappropriate given OP's plea of advice. In our data, offence was sparked in 23 out of 337 cases with pre-emptive clarifications (see Table 1). Although this constitutes a small percentage (6.8%), it illustrates how clarifications that ideally would clear any misunderstandings can be seen as inadequate defences.

5.1.2 Reparative claims

As posited above, the reparative function of 'only joking' claims is relevant after an attempt at humour has been misconstrued or criticised. Such claims of intent can be successful if they lead to the offended party withdrawing their criticism. However, in case of offence, the defence is often unsuccessful: despite the claim and recognition of a jocular intent, the joke may be deemed to be inappropriate. Example (6), from the subreddit r/India, is from a sprawling discussion on trains in India; the full thread has around 400 comments, but in one specific exchange, User A expresses mock incredulity at North-East India having train infrastructure. This short comment has received more than 200 upvotes, suggesting that readers generally recognise that the comment ('you guys have trains') is not a genuine question and may even be a reference to a popular meme template (*You guys are getting paid?* (n.d.) on KnowYourMeme). However, User B interprets User A as being serious, causing A to clarify their jocular intent ("chill bro, it's just a joke").

(6) A: You guys have trains? (+239)
B: People should really be updated in order to pass such comments. There are regular

trains to Arunachal Pradesh, Tripura, and Nagaland. (+37)

A: Chill bro its just a joke […] Also are you hurt, but i am the one who is hurt inside, in my entire existence i have never been to train. Our state got one recently but its at the border and not sure if passenger travelling is even started or not. (-8)

C: then crack some better jokes (+3)

A: Sorry if that offends you, are you from northeast? If not, thats the reason for not getting the joke. (0)

C: Give it a rest (+1)

Despite the reparative claim, the joke is deemed to be in bad taste by User C ("then crack some better jokes"). That is, although User A's jocular intent is acknowledged, they are held accountable for causing offence with an inappropriate or unfunny joke (Chang & Haugh 2021). User A offers an apology for causing offence, yet they also suggest that User C has misconstrued the joke about North-East India as offensive because they are not personally familiar with train infrastructure in North-East India. The sincerity of User A accepting accountability through an apology is thus undermined by the simultaneous sanctioning of User C for taking the joke too seriously. User A thus assigns accountability to User C but is downvoted for doing so. The argument eventually ends as C offers another metapragmatic act, with "give it a rest" functioning as an intervention, here produced by one of the debaters (see Tanskanen 2021).

A joke may be deemed offensive if it is inappropriate to the situation. In example (7), similarly to example (5), offense is sparked by an attempt at humour in a situation that other interactants see as a serious exchange of giving advice. User A has asked for career advice as a highly educated engineering professional, but User B suggests that User A work as an intern, pre-emptively marking a jocular tone with "lol jk".

(7) A (OP): Pretty soon here I'll be a licensed PE [=Professional engineer]. But I'm wondering what sort of role I could be a good match for if I ever went to a consulting firm, as I haven't even touched CAD [computer-aided design] since college. What are your thoughts?

B: Starting over with the interns lol jk maybe you could push some paper (-2)

A: Maybe they'll even let me answer the phones (+2)

B: Ya I mean I dunno what you expect. We do CAD all day lol (-3)

A: Well since I haven't worked at a firm I didn't know what to expect. What I didn't expect was a snarky ass response when I was showing general curiosity. (+2)

B: Bruh it's the truth lol you don't know how to design lol I don't know what you expect […] (-3)

At first, User A at first does not indicate taking offence: their second turn in (6) can be construed as going along with the non-serious frame. However, later they criticize User B's joke for being a "snarky ass response" rather than helpful advice; this turn implicates that they are expecting a serious response and rejects the invitation to mutual affiliation (see Haugh 2016). Yet the offending party maintains that their teasing doubles as truthful advice on career requirements ("it's the truth") and the repeated uses of clause-final *lol* likely also mark non-serious or non-aggressive illocutionary force (see Schneebeli 2020: 14). Whether or not User A accepts User B's claim of jocular intent, A appears offended by the non-serious answer. Ultimately, the discussion remains unresolved as A stops responding to B and neither party accepts accountability for taking or causing offence. Interestingly, User B's non-seriousness at A's expense is judged as

inappropriate in this advice-giving context also by the community because B's offending comments have been downvoted whereas comments by A have been upvoted. This suggests that onlookers in the community believe that User B, irrespective of their claimed intent, should be held morally accountable for causing offense through a belittling joke at User A's expense and withholding expressions of remorse.

5.2 Negotiating the joking intent of another user

The examples so far have illustrated cases where users clarify their own intent, whether pre-emptively or reparatively. However, in the polyloguous context of Reddit, users often clarify the intentionality of a joke (originally) made by someone else (192 cases, see Table 1). Example (8) illustrates a successful clarification of another user's jocular intent. In this case, the joke about a fictional character is originally made on the micro blog platform Tumblr, but by sharing a screenshot on Reddit, User A recontextualises it for the Reddit audience.

(8) [User A has posted a screenshot of a Tumblr post that discusses the television series *House of the Dragon*, focusing on the fictional character Helaena's deep interest in insects.]
B: tsk. Imagine you are a bigot enough to other an innocent little girl because she has different hobbies than you. That person would have had a stroke when you show them an entomologist, perhaps. (-2)
A: it's a joke (+4)
B: Really? 😳 Gone way over my head then. Sorry. (-7)
A: No problem, happens all the time on the internet. But yeah, from what I've seen, people on Tumblr love Helaena […] they're just making some jokes about her and her siblings lol (+6)

The reparative jocular claim comes after User B's reaction has treated the shared content as serious; claims following reproaches from others were recurrent in Haugh and Chang's (2021) spoken data as well. In (8), however, the jocular intent has been clear to others, as they have downvoted B's comment and upvoted A's clarification, signalling that they have the shared knowledge to understand that the intent was to amuse rather than offend (see Dynel 2021a; Yus 2021). The negotiation continues with User B's apology for misinterpreting the illocutionary force and acceptance of the jocular claim. User A acknowledges the apology before explaining the origin of the joke; A focuses on the intent of "people on Tumblr" and their ways of joking, taking the role of a mediator here, and, at the same time shifting accountability to them. But while the joke User A is defending is not their own, they are responsible for posting it and thus partially accountable for the intent(s) ascribed. Again, other members of the subreddit show their appreciation of User A's comment with upvotes. The sequence is similar to that of example (1) and shows reflexive accountability by both users as they negotiate towards a joint meaning.

A claim following a reproach is illustrated by example (9) as well, in which User B defends a post by another Redditor and stresses its jocular intent. The post in question is a copy of a tweet, which refers to ADHD as diet autism. The tweet ends with "lmfao", indicating that the poster views it as a joke. Other users acknowledge the jocular intent but consider the joke inappropriate. There seem to be two competing interpretations of the intent of the post/tweet: one seeing it as just jocular, the other perceiving it as both jocular and offensive.

(9) A: I actually find it really annoying to call ADHD diet autism. ADHD is it's own thing that can be comorbid with ASD. It is not a less severe form of autism. ADHD is ADHD. (+22)
B: Its just a joke (-11)
A: I know but I see it enough that I'm tired of it. (+8)
C: that much is obvious. People are trying to explain that despite it being a joke, it's not an okay joke to make. saying "it's just a joke" is a cheap way to try and invalidate how you offend or hurt people with your "jokes" (+3)

User A's explanation of why they find the post (tweet) annoying has received a good number of upvotes, signalling that other users appreciate and support the view expressed by A. User B's claim in defence of the tweet ("Its just a joke"), in contrast, has been downvoted, indicating rejection of the claim. The two further posts, both supported with upvotes, show that the jocular intent of the post is acknowledged but not accepted, and therefore, making a claim of jocular intent does not help in this situation. The posters have taken offence, and despite the recognised jocular intent, ascribe accountability for the intent to offend (see Dynel 2021a).

Similar competing interpretations of another Redditor's intent are evident in examples from one particularly fruitful thread on a subreddit devoted for sharing jokes and humorous memes. The stimulus for the discussion is a sexist meme, portraying a group of women going over technical-looking instructions on how to make a sandwich. The meme has been around for quite some time, but this does not stop Redditors from discussing it; there are nearly two thousand comments posted in the discussion. There seem to be (at least) three interpretations at play. Some users metapragmatically label the meme as "just a joke" (examples 10 and 11):

(10) It's just a joke, chill. No one is forcing you to find everything funny (+2)
(11) It's funny because people are butthurt, that is actually funny. If you can't take a joke, you need to reevaluate your own self esteem. (+8)

For the posters of (10) and (11), the jocular intent of the meme excuses any potentially sexist or offensive content, although what is deemed funny varies between the posters: in (10), it is the meme itself, while in (11) it is the reaction caused by the meme. For another group of users, the meme is clearly sexist and offensive, and claims of jocular intent are rejected ("Not funny in the slightest. Very offensive."). The third interpretation focuses on the fact that the meme is not new. For these users, jocular intent might have been acceptable earlier, whereas now they find the meme to be "in bad taste", or simply too old to be either funny or offensive.

Based on the number of up- and downvotes, although only some posts have received a noticeably high score, all three major interpretations receive at least some support and there seems to be no genuine aim towards consensus. This exemplifies, on the one hand, successful communication where many participants agree on the quality and appropriateness of the joke. On the other hand, it also suggests that the audience is heterogenous. The subreddit in question is one of the most popular on the platform and it is devoted to such broad topic categories as "funny" and "memes" (as per the subreddit description). Perhaps because of this, it does not have a single consensus on offensiveness or inappropriateness, because the audience simply has too varying views on humour and offence to have a shared emic definition (Dynel & Poppi 2019).

In contrast, a less diverse audience may form clearer criteria and consensus for humour. A community formed around a shared ideology or goal may use humour and memes to create affinity between community members. Such communities may have such a local culture and specific understanding of humour that even as community-internal interaction is successful,

outsiders misinterpret the intention. The negotiation in example (12) is in response to a post about transgender people. Similar to example (9), User A recognises the jocular intent but sees the joke as perpetuating harmful stereotypes. However, this criticism is heavily downvoted and has also received verbal criticism and ridicule from User B, which in turn is upvoted.

(12) A: Actually we should be careful about what jokes we tell because too often we absorb their message uncritically and play out those learned lessons against real people. (-54)
B: LMAO TRIGGGEEEERRRRED ARRR SLASH THELEFTCANTMEME AMIRITE 🤣🤣🤣🤣🤣🤣🤣🤣🤣🤣🤣🤣 TRYNA CANCEL COMEDY NOW ARE WE YOU CANT TAKE A JOKE AHAHAHAHAHAHAHAHAHAHAAHAHA (+ 36)

The exchange cannot be taken at face value and contextual understanding of the subreddit norms is necessary. The description of the subreddit, r/transphobiacirclejerk, explicitly states "Parody subreddit for trans people, mocking all transgender-related topics. If you take this seriously, then don't." Many subreddits use parody of outgroup as a social critique, where parodying intent may be evident in hyperboles or sarcasm (Dynel 2020; Mullan 2022): here User B is mimicking right-wing discourse through references to cancel culture and the right-wing subreddit r/TheLeftCantMeme. However, User A ignores or does not recognise the norms of r/transphobiacirclejerk – offence is triggered because A is an unratified outsider to the in-group's performative ritual, unable to recognize that User B's discourse resources aim to mock transphobes as part of a conventionalized ritual practice (see Kádár 2013: 56–59). Meanwhile, in-group members recognize and approve of the jocular intent and the up- and downvotes suggest a consensus and general agreement of that it is locally acceptable and non-offensive to parody transphobic humour, ideologies and speech patterns (as perceived by this subreddit). This illustrates how subreddit rules and target audience should be accounted for both by Redditors and researchers evaluating the offensiveness of the joke.

5.3 Intent and editability of Reddit

The negotiation of intent is given another layer of complexity by some of the technological features of Reddit that afford users CMC-specific means to mark intent. In this section we consider negotiation of intent in a context where contributions can be edited. Edited content is challenging to retrieve systematically as our metadata did not include multiple versions of comments. However, we observed some instances in our data where edits were pointed out in the comment text. In some cases such as (13), posts that had initially had no explicit claim of jocular intent had been subsequently edited by the users to clarify the intent reparatively. In (13), User A initially jokes about the audience vote of the Eurovision Song Contest, which sparks verbal criticism from User B as well as downvotes. Ten hours after B's complaint, User A responds by apologising to B and editing their original comment by adding a "/j" tag and apologising for their unclear intent. The original joke remains in its original form, and the subsequently added content is clearly marked in as "Edit:", a customary practice.

(13) A [Original 2:03 AM[4], Edited 12:03 PM]: Australia only getting 2 points from the televote was homophobic AND ableist, change my mind. Edit: /j lol I don't actually think people are ableist if they didn't vote for Not The Same. Sorry it wasn't clear. (-11)
B [5:58 AM]: girl what? how could anyone possibly know that on the night of the final.

[4] Time stamps retrieved from Reddit as UTC. Local time of the users in unknown.

most people that watch eurovision dont know anything except what they see on the night. if the audience were that ableist then salvador wouldn't have won with the things people thought about him. (+15)
A [12:06 PM]: Sorry; my comment was a joke that didn't land well. Ofc you're not actually ableist or homophobic if you didn't like Australia. (+1)

However, comments can also be edited so that the content of the joke itself is altered to be less likely to cause offense. In example (14), the discussion starts with User A's reaction to a recent LCK esports match: "the longest 24-minute game" evaluates the match as slow-paced. In response, User B implies that the pace was fast by esports standards, prefacing this with "unironically" (14). Confusion over User B's intent may stem from the metapragmatic *unironically* as well as the lack of overt exaggeration or objectively untruthful claims that could signal the statement being a jocular insult against the LCK team (see Dynel 2021a). Some users, including User C, criticise B's comment; 3.5 hours later, B edits their comment by removing "unironically" and adds a call for other users to "take a joke".[5] User B thus reproaches users who had interpreted the comment as serious.

(14) A [1:45 AM]: That was the longest 24-minute game I've ever seen.
B [Original 1:52 AM]: Unironically most fast-paced LCK team
C [5:04 AM]: Man, that's not even true. Sometimes I think you people don't even watch LCK.
B [Edited 5:20 AM]: Most fast paced LCK team. Edit: LCK frogs in the replies take a joke

Note that in (14), User B's label "Edit:" does not refer to the removal of the word "unironically", and this change is not visible to readers on the forum. This undisclosed edit was brought to our attention by another user, who appears familiar with the original comment and criticises the edit for dishonest intention (15).

(15) D [3:33 PM]: Lmao sick edit bro. First started the comment with "un-ironically" and then you edit that out and say it was a joke? Either you don't know what un-ironically means or you are dishonest.

Examples (14) and (15) show how the editability of Reddit posts complicates the interpretation of intentionality – both for other users and for researchers. One cannot infer whether B's reparative "take a joke" indicates a genuine jocular intent or merely claims jocular intent to defend against negative attention. Editability affords communicators control over the content and context of their message even after it has been received (see Treem & Leonardi 2013: 159–160). That is, not only are Reddit users able to state their intent *post facto* in an effort to repair unsuccessful communication, but they may also seek to correct the content of their offensive message even after it has already been viewed. A comprehensive analysis of editability is beyond the scope of this paper, but it would be interesting to further examine this affordance as a reparative strategy.

6 Discussion & conclusion

This paper has explored how users on the social media platform Reddit negotiate instances of offensive humour. The negotiation of offense and jocular intent seeks to prevent or solve a

[5] The edited post was taken from our dataset and was attested on Reddit.com during analysis, but the original pre-edit text was retrieved from a record in the Internet Archive, https://web.archive.org.

misunderstanding of the user's intent. In order to locate such negotiations, we started with metapragmatic claims of jocular intent, with which users comment on and clarify the intent of a post. Other users can voice their interpretations and explain them, and divergent meanings can be further reviewed in the ensuing metapragmatic negotiation. At the same time, accountability for both the intent and possible inferences is ascribed and negotiated. Many of our examples illustrated how users can arrive at a joint understanding and achieve successful interaction even in cases where the writer's claimed intent and the recipient's interpretation did not originally match (cf. Elder & Haugh 2018; Elder 2021b).

The findings of our analysis of online interaction show similarities to those presented for spoken interaction as regards the functions of jocular claims (see Chang & Haugh 2021). Redditors use them as pre-emptive disclaimers and reparative claims, both in cases where the offensive joke is defended by its instigator and where the joke is created by someone else. Some of the claims that follow a reproach sanction another user for taking the joke too seriously. Yet others signal a transition from humorous to serious. The analysis also showed that all types of claims can be successful or unsuccessful in clarifying humour or repairing failed humour. Furthermore, even when the jocular intent is acknowledged, the offending party may be held accountable for causing offence through criticism of their sense of humour or lack of tact and consideration, as in our example on ADHD (9).

Whereas individual users' responses and reactions form a *micro-context* for interpreting intent, Reddit and subreddit norms provide a *macro-context* for the interpretation of humour, offence, and appropriateness of content (Dynel 2021a). Upvotes and downvotes, part of the micro-context, allow participants themselves to weigh in non-verbally in whether a piece of content is permissible given the local context; this information provides researchers context for assessing overall reactions to both sides in a negotiation of offence and accountability. In our data, offense is clearly linked to macro-context community norms of appropriateness.

The success of a jocular claim is decided locally by the users in the specific interactional context of the subreddit (Dynel 2023: 242). Furthermore, the negotiation of humour frequently represents boundary work (Bell 2015: 156; Shifman 2019: 47). Circulating or defending humour may align a user with a community whereas criticising or downvoting inappropriate humour is a way for Reddit users to regulate group norms and boundaries. Similarly, criticising a user for being offended aims to draw boundaries that exclude the offended party as an outgroup member. Boundary work is attested when users fail to recognise the jocular intent (examples (6), (7) and (8)) or when they fail to recognize the "licence for offensive humour", even when it is provided by explicit community rules on parody (example (12)). Indeed, users defending the jocular intent of another person do so because they are privy to the in-group norms, rituals and intertextual references whereas the offended individual is not. The metapragmatic judgements of humour may thus draw on collective identities to explain why something is locally acceptable or offensive (see Sinkeviciute 2019). Yet norms and group boundaries are not always clear-cut. In the case of the sandwich meme (examples 10 and 11), licence to both cause and take offence are evident, because the heterogeneous audience has clearly diverse opinions about the joke; there are competing interpretations, and no clear consensus.

By searching for metapragmatic claims relating to jocular intent, we have focused on the negotiation of offense rather than on the (potentially) offensive content itself. We acknowledge that other ways of negotiating offence and jocular intent are outside the scope of our approach. However, the corpus pragmatic approach allows for searching for an infrequent yet widespread

metapragmatic act. Conducting the search in a dataset across the whole of Reddit emphasises that the need to negotiate offensive humour is not restricted by the topic of the discussion: jocular claims are encountered on subreddits dealing with a great variety of themes and purposes. This attested range emphasises the need to consider offense and its negotiation also in contexts that at first glance hardly seem divisive. Our analysis has demonstrated that offense can be sparked by jokes on sensitive topics but also by situations where users misread the context or have different expectations about appropriate conduct as is common in diverse userbases (see Tagg, Seargeant & Brown 2017). That different online and offline settings vary in their norms further emphasises the need for an emic understanding of negotiation of offence in interaction.

References

Allison, Kimberley R., Kay Bussey & Naomi Sweller. 2019. "I'm going to hell for laughing at this": Norms, humour, and the neutralisation of aggression in online communities. *Proceedings of the ACM on human-computer interaction* 3(CSCW). 1–25. https://doi.org/10.1145/3359254

Altahmazi, Thulfiqar H. 2024. Humorous but hateful: Linguistic impoliteness and visual dysphemism in anti-Muslim memes. *Internet Pragmatics* 7. 219–248. https://doi.org/10.1075/ip.00106.alt

Andersson, Marta. 2021. The climate of climate change: Impoliteness as a hallmark of homophily in YouTube comment threads on Greta Thunberg's environmental activism. *Journal of Pragmatics* 178. 93–107. https://doi.org/10.1016/j.pragma.2021.03.003

Arendholz, Jenny. 2013. *(In)Appropriate online behaviour: A pragmatic analysis of message board relations*. Amsterdam: John Benjamins.

Aslan, Erhan. 2022. Days of our 'quarantined' lives: Multimodal humour in COVID-19 internet memes. *Internet Pragmatics* 5. 227-256. https://doi.org/10.1075/ip.00075.asl

Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire & Jeremy Blackburn. 2020. The Pushshift Reddit dataset. *Proceedings of the International AAAI Conference on Web and Social Media* 14. 830–839. https://doi.org/10.1609/icwsm.v14i1.7347

Bell, Nancy D. 2015. *We are not amused: Failed humor in interaction*. Berlin: Mouton de Gruyter.

Blommaert, Jan & Ben Rampton. 2011. Language and superdiversity. *Diversities* 13. 1–22.

Bou-Franch, Patricia. 2021. Evaluation, conflict and prescriptive metapragmatic comments: (re)constructing transmedia stories. In Marjut Johansson, Sanna-Kaisa Tanskanen & Jan Chovanec (eds.), *Analyzing digital discourses: Between convergence and controversy*, 189–217. Cham: Palgrave Macmillan.

Caffi, Claudia. 1998. Metapragmatics. In Jacob L. Mey (ed.), *Concise encyclopedia of pragmatics*, 581–586. Amsterdam: Elsevier.

Chang, Wei-Lin Melody & Michael Haugh. 2021. Teasing and claims to non-serious intent in Chinese talk shows. *East Asian Pragmatics* 6. 135–159. https://doi.org/10.1558/eap.18158

Culpeper, Jonathan. 2011. *Impoliteness: Using language to cause offence*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511975752

Demjén, Zsófia. 2016. Laughing at cancer: Humour, empowerment, solidarity and coping online. *Journal of Pragmatics* 101. 18–30. https://doi.org/10.1016/j.pragma.2016.05.010

Dynel, Marta. 2008. No aggression, only teasing: The pragmatics of teasing and banter. *Lodz papers in pragmatics* 4(2), 241–261.

Dynel, Marta. 2017. Is there a tumour in your humour? On misunderstanding and miscommunication in conversational humour. In Rachel Giora & Michael Haugh (eds.), *Doing Intercultural Pragmatics: Cognitive, Linguistic and Sociopragmatic Perspectives on Language Use*, 55-78. Berlin: Mouton de Gruyter.

Dynel, Marta. 2020. Vigilante disparaging humour at r/IncelTears: Humour as critique of incel ideology. *Language & Communication* 74. 1–14. https://doi.org/10.1016/j.langcom.2020.05.001

Dynel, Marta. 2021a. Desperately seeking intentions: Genuine and jocular insults on social media. *Journal of Pragmatics* 179. 26–36. https://doi.org/10.1016/j.pragma.2021.04.017

Dynel, Marta. 2021b. Humour and (mock) aggression: Distinguishing cyberbullying from roasting. *Language & Communication* 81. 17–36. https://doi.org/10.1016/j.langcom.2021.08.001

Dynel, Marta & Fabio I.M. Poppi. 2019. Risum teneatis, amici?☆: The socio-pragmatics of RoastMe humour. *Journal of Pragmatics* 139. 1–21. https://doi.org/10.1016/j.pragma.2018.10.010

Dynel, Marta. 2023. Irony and humor. In Raymond W. Gibbs, Jr & Herbert L. Colston (eds.), *The Cambridge Handbook of Irony and Thought*. Cambridge: Cambridge University Press, 237-254. doi:10.1017/9781108974004.019

Dynel, Marta. 2024. Do we know whether to laugh or cry? User responses to @Ukraine's dark-humour meme. *Journal of Creative Communications* 19(3), 243-258. https://doi.org/10.1177/09732586241239908

Elder, Chi-Hé. 2021a. Microaggression or misunderstanding? Implicatures, inferences and accountability. *Journal of Pragmatics* 179. 37–43. https://doi.org/10.1016/j.pragma.2021.04.020

Elder, Chi-Hé. 2021b. Speaker meaning, commitment and accountability. In Michael Haugh, Daniel Kádár & Marina Terkourafi (eds.), *The Cambridge handbook of sociopragmatics*, 48–68. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108954105.005.

Elder, Chi-Hé & Michael Haugh. 2018. The interactional achievement of speaker meaning: Toward a formal account of conversational inference. *Intercultural Pragmatics* 15. 593–625. https://doi.org/10.1515/ip-2018-0021

Elder, Chi-Hé & Michael Haugh. 2023. Exposing and avoiding unwanted inferences in conversational interaction. *Journal of Pragmatics* 218. 115–132. https://doi.org/10.1016/j.pragma.2023.09.014

Erjavec, Karmen & Melita Poler Kovačič. 2012. "You don't understand, this is a new war!" Analysis of hate speech in news web sites' comments. *Mass Communication and Society* 15. 899–920. https://doi.org/10.1080/15205436.2011.619679

Garfinkel, Harold. 1984(1967). *Studies in Ethnomethodology*. Cambridge: Polity.

Glenski, Maria, Greg Stoddard, Paul Resnick & Tim Weninger. 2018. GuessTheKarma: A game to assess social rating systems. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW). 1–15. https://doi.org/10.1145/3274328

Grice, Herbert Paul. 1957. Meaning. *Philosophical Review* 66. 377–388.

Haddington, Pentti. 2011. Serious or non-serious? Sequential ambiguity and disavowing a prior stance. *Functions of Language* 18. 149–182.

Hansen, Maj-Britt Mosegaard & Marina Terkourafi. 2023. We need to talk about Hearer's Meaning! *Journal of Pragmatics* 208. 99–114. https://doi.org/10.1016/j.pragma.2023.02.015

Haugh, Michael. 2007. The co-constitution of politeness implicature in conversation. *Journal of Pragmatics* 39. 84–110. https://doi.org/10.1016/j.pragma.2006.07.004

Haugh, Michael. 2008. Intention in pragmatics. *Intercultural Pragmatics* 5. 99–110. https://doi.org/10.1515/IP.2008.006
Haugh, Michael. 2015. Impoliteness and taking offence in initial interactions. *Journal of Pragmatics* 86. 36–42. https://doi.org/10.1016/j.pragma.2015.05.018
Haugh, Michael. 2016. "Just kidding": Teasing and claims to non-serious intent. *Journal of Pragmatics* 95. 120–136. https://doi.org/10.1016/j.pragma.2015.12.004
Haugh, Michael. 2018. Corpus-based metapragmatics. In Andreas H. Jucker, Klaus P. Schneider & Wolfram Bublitz (eds.), *Methods in pragmatics*, 619–644. Berlin etc.: De Gruyter Mouton.
Hayes, Rebecca A., Caleb T. Carr & Donghee Yvette Wohn. 2016. One click, many meanings: Interpreting paralinguistic digital affordances in social media. *Journal of Broadcasting & Electronic Media* 60(1). 171–187. https://doi.org/10.1080/08838151.2015.1127248
Hübler, Axel & Wolfram Bublitz. 2007. Introducing metapragmatics in use. In Wolfram Bublitz & Axel Hübler (eds.), *Metapragmatics in use*, 1–26. Amsterdam & Philadelphia: John Benjamins.
Kádár, Daniel. 2013. *Relational rituals and communication: Ritual interaction in groups*. London: Palgrave Macmillan. https://doi.org/10.1057/9780230393059
Kleinke, Sonja & Birte Bös. 2015. Intergroup rudeness and the metapragmatics of its negotiation in online discussion fora. In Miriam Locher, Brook Bolander & Nicole Höhn (eds.), *Relational work in CMC. Special issue of Pragmatics* 25. 47–71.
Kohnen, Thomas. 2007. Text types and the methodology of diachronic speech act analysis. In Susan M. Fitzmaurice & Irma Taavitsainen (eds.), *Methods in historical pragmatics*, 139–166. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197822.139
Lave, Jean & Etienne Wenger. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
Liu, Ping & Huiying Liu. 2021. Salience adjusting: Metapragmatic expressions in complaint responses. *Journal of Pragmatics* 176. 150–163.
Marwick, Alice E. & danah boyd. 2011. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society* 13(1). 114–133. https://doi.org/10.1177/1461444810365313
Massanari, Adrienne L. 2017. #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society* 19(3). 329–346. https://doi.org/10.1177/1461444815608807
Massanari, Adrienne L. 2019. "Come for the period comics. Stay for the cultural awareness": reclaiming the troll identity through feminist humor on Reddit's /r/TrollXChromosomes. *Feminist Media Studies* 19(1). 19–37. https://doi.org/10.1080/14680777.2017.1414863
Mitchell, Nathaniel & Michael Haugh. 2015. Agency, accountability and evaluations of impoliteness. *Journal of Politeness Research* 11. 207–38. https://doi.org/10.1515/pr-2015-0009
Mullan, Kerry. 2022. On the "Dark Side": Facebook humour used for inclusion and exclusion. *The European Journal of Humour Research* 10. 96–115. https://doi.org/10.7592/EJHR2022.10.2.644
O'Keeffe, Anne. 2018. Corpus-based function-to-form approaches. In Andreas H. Jucker, Klaus P. Schneider & Wolfram Bublitz (eds.), *Methods in pragmatics*, 587–618. Berlin: De Gruyter. https://doi.org/10.1515/9783110424928-023
Orthaber, Sara. 2019. Aggressive humour as a means of voicing customer dissatisfaction and creating in-group identity. *Journal of Pragmatics* 152. 160–171. https://doi.org/10.1016/j.pragma.2018.11.002

Robards, Brady. 2018. Belonging and neo-tribalism on social media site Reddit. In Anne Hardy, Andy Bennett & Brady Robards (eds.), *Neo-tribes: Consumption, leisure and tourism*, 187–206. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-68207-5_12

Reddiquette. 2025. Reddit Help. Available online: https://support.reddithelp.com/hc/en-us/articles/205926439-Reddiquette (Accessed 12.6.2025)

Schegloff, Emanuel A. 2001. Getting serious: joke → serious 'no'. *Journal of Pragmatics* 33. 1947–1955. https://doi.org/10.1016/S0378-2166(00)00073-4

Schneebeli, Célia. Where *lol* Is: Function and position of *lol* used as a discourse marker in YouTube Comment. *Discours* 27. https://doi.org/10.4000/discours.10900

Shifman, Limor. 2019. Internet memes and the twofold articulation of values. In Mark Graham & William H. Dutton (eds.), *Society and the internet: How networks of information and communication are changing our lives* (2nd ed), 43–57. Oxford: Oxford University Press.

Sinkeviciute, Valeria. 2019. Juggling identities in interviews: the metapragmatics of 'doing humour'. *Journal of Pragmatics* 152. 216–227.

Skalicky, Stephen, Cynthia M. Berger & Nancy D. Bell. 2015. The functions of "just kidding" in American English. *Journal of Pragmatics* 85. 18–31. https://doi.org/10.1016/j.pragma.2015.05.024

Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Tagg, Caroline, Philip Seargeant & Amy Aisha Brown. 2017. *Taking offence on social media: Communication and conviviality online*. London: Palgrave.

Tanskanen, Sanna-Kaisa. 2007. Metapragmatic utterances in computer-mediated interaction. In Wolfram Bublitz & Axel Hübler (eds.), *Metapragmatics in use*, 87–106. Amsterdam & Philadelphia: John Benjamins.

Tanskanen, Sanna-Kaisa. 2018. Identity and metapragmatic acts in a student forum discussion thread. In Birte Bös, Sonja Kleinke, Sandra Mollin & Nuria Hernández (eds.), *The discursive construction of identities on- and offline*, 133–152. Amsterdam & Philadelphia: John Benjamins.

Tanskanen, Sanna-Kaisa. 2021. "Stop arguing": Interventions as metapragmatic acts in discussion forum interaction. In Marjut Johansson, Sanna-Kaisa Tanskanen & Jan Chovanec (eds.), *Analyzing digital discourses: Between convergence and controversy*, 219–244. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-84602-2_9

Tanskanen, Sanna-Kaisa. 2024. "Ding ding ding we have a hypocrite!": The metapragmatics of verbal hypocrisy in discussion forum interaction. In Sandrine Sorlin & Tuija Virtanen (eds.), *The pragmatics of hypocrisy*, 74–95. Amsterdam & Philadelphia: John Benjamins.

Treem, Jeffrey W. & Paul M. Leonardi. 2013. Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association, *Annals of the International Communication Association* 36. 143–189. https://doi.org/10.1080/23808985.2013.11679130

Verschueren, Jef. 2000. Notes on the role of metapragmatic awareness in language use. *Pragmatics* 10. 439–456.

You guys are getting paid? n.d. KnowYourMeme, https://knowyourmeme.com/memes/you-guys-are-getting-paid (Retrieved 12.6.2025)

Yus, Francisco. 2021. Pragmatics, humour and the internet. *Internet Pragmatics* 4. 1–11.

Zappavigna, Michele. 2022. Social media quotation practices and ambient affiliation: Weaponising ironic quotation for humorous ridicule in political discourse. *Journal of Pragmatics* 191. 98–112. https://doi.org/10.1016/j.pragma.2021.12.003