"We could shoot him…"
Interpretations of offensive humour

June 2025

Abstract

In everyday dialogue, people often say things that may be or are taken as offensive by their hearers. Whether or not these utterances are actually taken as offensive is highly context sensitive. Particularly important are social aspects such as the personae of the speaker and audience. Different groups of listeners have access to – and indeed embrace – different background assumptions and entertain different attitudes with regard to these. We consider such assumptions to be Aristotelian *topoi* and a set of topoi significant for a particular group in a particular context to be a *topoplex.* In this paper we present three real world examples of supposedly humorous and potentially offensive utterances. Our analysis is compatible with an established theory of dialogue semantics which formalises topoi as mechanisms for common sense reasoning that arise from specific interactional experiences.

1 Introduction

In everyday dialogue, people often say things that may be or are taken as offensive by their hearers. In a similar, but orthogonal vein, people may say things that their hearers take to be humorous – regardless of whether or not either offence or humour was the intention of the speaker.[1] These factors interact in the case of *offensive humour*, where people may understand something someone else says to be offensive but nevertheless find it humorous.

Whether potentially offensive utterances are taken as being offensive depends on a large number of factors, including, but not limited to, the subject of the utterance, the context of the utterance, the identity of the speaker with respect to the potentially offensive topic, the identity of the hearer with respect to the potentially offensive topic and the relationship between speaker and hearer (Culpeper 2011: Chapter 6). As can be seen here, many of these factors rely on

---

[1]Note that our goal here is not to disambiguate utterances humorous from non-humorous (Warren & McGraw 2016) or those that can be taken as offensive or not, as we believe that both humour and offensiveness are highly subjective. We focus on the features of the context which give rise to such subjective judgements/experiences and remain agnostic about whether such categorisation is even possible.

the social and identity characteristics of speaker and listener, and these factors of social context are, we argue, critical and often overlooked in discussions of offence in general and offensive humour specifically.

As speakers, in order to achieve a humorous effect from a potentially offensive utterance, we have to judge both the level of offence the audience are likely to take (too much is simply offensive) and whether they share certain background assumptions. In a closed setting, such as in a casual conversation between friends, this is usually unproblematic. However, in our increasingly hyper-connected world, the recipient of the offence is underspecified, meaning that this is a non-trivial issue, which, as we shall see, cannot be fully determined. We are therefore motivated by the need for a comparative study of offensive utterances marked as humorous by public speakers, which have caused different degrees of offence in different audiences.

In this paper we will examine three examples which different audiences rated as offensive and/or humorous. We will discuss how the different interpretations can be explained using existing mechanisms from an established theory of dialogue semantics (see section 2.1) which foregrounds the idea of mismatches between our personal takes on any interaction, and the common sense reasoning we rely on in conversation (Breitholtz 2020). Any comprehensive theory of dialogue needs to be able to account for such potentially problematic borderline cases involving public discourse with diverse audiences.

In section 2 we will outline related work and the background for our analysis. In section 3, we present our examples and discuss and analyse different reactions to them. Section 4 outlines steps for generalising our approach and section 5 provides a general discussion of the implications.

2 Background and related work

It is well known that context is essential to pragmatic interpretation, and many theories are devoted to its analysis (Van Dijk 2008, Linell 2009). For humour, although the situational context is important – what is appropriate to say at a funeral might not be appropriate at a wedding – social context arguably matters just as much, if not more, as humour often relies on inferences based on background knowledge, cultural assumptions and interpersonal experience rather than situational factors (Kuipers 2015, Davies 1990).

It has been shown that the kind of information that is used to warrant humorous inferences can be modelled as *topoi* – "soft" associative rules of thumb which underpin everyday inferences and are community specific (Breitholtz & Maraev 2019, Maraev et al. 2021). By "soft" we mean that they are assumed to often hold, or to hold in a typical case, but not by necessity. For example, it is a generally accepted topos that birds fly, licencing something like the inference "if x is a bird then x flies", which we represent as in (1).[2] Even though most of

---

[2]Following Breitholtz & Howes (2020), we use this semi-formal notation to represent our intuitions about enthymemes and topoi. The wiggly line denotes an associative rather than a strictly logical chain of reasoning.

us are aware that this is not universally true, given our knowledge of penguins (2)-(3) and ostriches, we might still draw on this topos in an argument such as "don't jump off that cliff – you are not a bird" in understanding that the warranted inference is that you, not being a bird, cannot fly.

$$\frac{\text{x is a bird}}{\text{x flies}} \tag{1}$$

$$\frac{\text{x is a penguin}}{\text{x cannot fly}} \tag{2}$$

$$\frac{\text{x is a penguin}}{\text{x is a bird}} \tag{3}$$

By "community specific" we mean that topoi are socially constituted within a community. Thus, topoi can differ between communities based on the experiences and interactions of the members of the communities (Rosengren 2002), which is a key component in many types of jokes (Davies 1998).

Topoi have been successfully used to model the basis for inferences in dialogue (Breitholtz 2020). Arguments put forward in everyday conversation are often *enthymematic* – i.e. the conclusion does not follow by necessity, usually because one or more premises are not explicitly stated. Recognising an acceptable topos that can underpin the enthymeme is crucial for understanding the dialogue (Breitholtz, Howes & Cooper 2023). In rhetoric and dialectic, topoi work as scaffolding for enthymemes (Aristotle, ca. 340 B.C.E./2007). Topoi can represent everything from facts acceptable to all inhabitants of planet earth (such as if we let go of something it generally falls downwards), to aspects of common ground that are specific to a subculture (e.g. that in the UK you need to order your drink at the bar rather than expecting table service), family norms (such as not having phones at the dinner table) or even just an understanding between two individuals.[3]

2.1 Reasoning in interaction

Much work on reasoning in interaction focuses on argumentation, political debate and media discourse targeting particular groups (Musolff 2004, Ball & Thompson 2017). While it is true that language whose essential purpose is influencing – and possibly deceiving – listeners relies heavily on considerations such as audience design (Bell 1984) and non-logical inference, such – essentially rhetorical – features are present in all types of interaction. For example, Jackson & Jacobs (1980) show that *enthymemes*, arguments that draw on implicit and sometimes questionable premises and generally considered a feature of rhetoric, are common in spontaneous conversation. Ducrot (1988, 1980) and

---

[3]Note that topoi are in general defeasible, although they may be more or less likely to hold in a given context – for example, if you let go of a helium balloon it floats upwards, or if you have a guest for dinner they may be allowed to use their phone at the table, but these rules of thumb are applicable in many situations and thus a useful cognitive resource.

Anscombre (1995) argue that topoi are essential not only for argumentation but for all kinds of interaction, as they supply implicit information and principles of reasoning which must be recognised by an interlocutor for enthymematic discourse to make sense. Consider, for example, the three conversations below between Arthur (aged 82) and his wife Evelyn (75).

(4) *From BNC KBB*
Arthur	77	Half inch of frost and you get the salt people out.
Evelyn	78	Mm.
	79	Well <unclear>.
Arthur	80	As soon as it's strikes your, the bottom of your car it's it begins to rust.
Evelyn	81	Well
Arthur	82	And ne , e it, and is, if it's going to freeze tonight you can't wash it off.

In order to understand this extract, you need to have the culturally specific knowledge that if the temperature falls below freezing then salt is put on the roads in the UK (by "the salt people"). You further need to link this knowledge to Arthur's following utterances, such that if there is salt on the roads and you drive on them, the salt strikes the underside of the car and causes it to rust, and that to prevent this rusting you need to wash the salt off. Additionally you need the general knowledge that the water you would use to wash the salt off would freeze if the temperature is below freezing, thus you can't wash it off. None of these are stated explicitly in the interaction.

(5) *Later the same day*
Evelyn	441	Now <pause> tomorrow <pause> we've got to get a sack of potatoes.
	442	We could wait till Monday but that's about all.
Arthur	443	We'll wait until Monday when there's salt on the roads.
Evelyn	444	Why?
Arthur	445	Well you know what salt does to underneath the car.
	446	[The bodywork].
Evelyn	447	[Well] I know!
	448	You're always on about that!

In the second dialogue Evelyn states the need for them to buy more potatoes by Monday at the latest as they are running out. Arthur comments that they will wait until Monday, for which utterance Evelyn cannot find or accommodate a suitable topos, prompting her "why?". Arthur's response, in line 445, relies on the same topos as in the previous interaction, namely that if there is salt on the roads and you drive on them, the salt strikes the underside of the car and causes it to rust, which Evelyn is now able to access. This leads to conflicting conclusions, and an interpretation that Arthur's utterance in line 443 must have been sarcastic or non-serious, because salt being on the roads and the damage

he claims it does to the car is a reason for Arthur not to (want to) drive (and not a reason for his wanting to drive).[4] This can be thought of as a very general topos saying that if x says p and you can infer not p – either because it is clear from the context that p is not true (such as telling someone who is obviously failing at something that they seem to be doing well), or, as in this case, based on common background assumptions – then you can infer that x is not being serious.

In the third dialogue, taking place the following day, the couple are again discussing the scarcity of remaining potatoes.

(6)　*The following day*

Arthur	1709	[What about the potatoes?]
Evelyn	1710	Well there's not many left.
Arthur	1711	Oh well, they'll be alright.
Evelyn	1712	That's what I say, we've go , we've got to go <pause> first time out next week
Arthur	1713	Mm.
Evelyn	1714	ge go to the farm and get a bag.
Arthur	1715	Well I'm not going out when the salts on!
Evelyn	1716	I know!
	1717	I know, keep the car
Arthur	1718	Well it's not that, [I've gotta]
Evelyn	1719	[for posterity!]
Arthur	1720	I've gotta wash it off haven't I?

Here Arthur is once again adamant they don't take the car out while the salt is on (making explicit what had to be inferred the previous day) and Evelyn is making a sarcastic comment about his reasons for this, drawing on the topos that Arthur provided in their previous interaction.

In addition to needing the culturally specific background knowledge that in the UK salt is put on the roads in below freezing temperatures, there are some very specific inferences that Evelyn and Arthur are able to make because of their (extensive) shared conversational history. The enthymematic argument conveyed is (among others) the following:

$$\frac{\text{there is salt on the road}}{\text{Arthur will not drive on the road}} \quad (7)$$

This enthymeme draws on principles of reasoning that are accepted, or at least acceptable, in the context. Salient topoi here have to do with general knowledge, cultural knowledge and participant specific knowledge:

---

[4]Here it is also relevant to note that several other topoi are required for Evelyn and Arthur to make sense of the interaction, such as couple specific topoi underpinning their acceptance that "getting potatoes" requires driving (perhaps the shop or farm is too far to walk, or the potatoes would be too heavy to carry, or is in a rural area lacking public transport options), but we leave these aside for now.

$$\frac{\text{temperature is below zero}}{\text{there is salt put on the roads}} \tag{8}$$

$$\frac{\text{salt is on the roads} \quad \text{drive on the roads}}{\text{the bodywork of the car gets ruined}} \tag{9}$$

$$\frac{\text{salt is on the roads} \quad \text{drive on the roads}}{\text{salt needs to be washed off the car}} \tag{10}$$

Indeed Arthur's somewhat elliptical utterances rely on the knowledge that he and Evelyn have talked about the salt on the roads before, but from two different perspectives – about the effect on the car (4:80) and from the perspective of the personal consequences (4:82). These two different perspectives require different topoi to warrant them, and while Evelyn appears to draw on the underpinning warrant (9) for the effect on the car (in lines 1717-1719), in this case in line 1720 Arthur draws on the alternative topos (10).

Arthur is further drawing on a topos mentioned in the conversation in (4), which is not explicitly reraised here but underpins his reluctance to drive the car on the roads if there is salt on them, as he can infer that it is not just that it would be effortful and time-consuming for him to have to wash the salt off the car, but is actually not even possible.

$$\frac{\text{temperature is below zero}}{\text{salt cannot be washed off the car}} \tag{11}$$

As shown in Breitholtz, Howes & Cooper (2023), interpretation and production of utterances is driven, to some extent, by the sets of topoi one entertains – or is willing to accommodate (Lewis 1979, Beaver & Zeevat 2007, Stalnaker 2014). As in the case of presupposition accommodation, accommodated topoi are often already part of background knowledge. However, in many cases topoi are constructed on the fly based on utterances in an interaction, and tentatively adopted in the dialogue. For example, Breitholtz & Maraev (2019) show how joke-telling sometimes requires that the dialogue participants adopt a topos which is clearly not part of background knowledge, like "if you want to put an elephant in a fridge you open the door, put the elephant in and close the door". Whether or not arguments in discourse are seen as valid depends on the topoi one accepts, and misunderstandings or disagreements in dialogue may be related to a mismatch of topoi between speakers. The ability to follow various strains of reasoning – including inconsistent ones – seems to be a prerequisite for the complex type of interactive language understanding and problem solving that humans master so well.

Standard mechanisms for evoking and accommodating topoi in interaction, described formally in Breitholtz (2020), can be exploited in several different contexts – both benign ones and those that are potentially less so.

2.2   Social context and topoi

As we have seen, topoi that are relevant to an interpretation might represent generally accepted norms or facts, or they might be previously established between the speakers or in a community to which the speakers belong. They can be derived from general knowledge, general or situation specific socio-cultural norms and experience of interactions with specific people, or be created on the fly in the course of a conversation. They are not restricted to the semantic content of an utterance (see for example Noble, Breitholtz & Cooper 2020), and can underpin inferences we make based on the social context. Social context can be couched in different ways, for example in terms of social networks (Milroy & Milroy 1978, Lev-Ari 2018) or in terms of the personae of participants in the interaction. In the game theoretic accounts of Henderson & McCready (2018) and Burnett (2020) personae are modelled as pairs of personality traits associated with attitudes with respect to a particular issue that contribute different social meaning interpretations to utterances. In Henderson and McCready's work, personae are used to model dogwhistle messages, that is, utterances that are interpreted in one way by one part of an audience while sending a different, and often more controversial message to a sub part of the audience.

In the context of humour the sender of an intentionally humorous message does not typically intend for only part of the audience to appreciate the joke, rather the speaker has a general idea of a larger set of topoi that are known or acceptable to his or her audience. According to Noble, Breitholtz & Cooper (2020), if a particular persona is projected to an individual, it evokes a set of topoi (which we refer to as a topoplex) associated with probability distributions, where the probability corresponds to how likely it is that this topos is evoked. If the probability is too low then the topos needs some extra information in order to be accommodated, i.e. marked as humorous or sarcastic, and if it is not understood as being such it might give rise to conversational repair (Mazzocconi, Maraev & Ginzburg 2018).

2.3   Offensive humour

Existing research recognises the importance of different forms of offence in everyday dialogue. Offence has been associated with such notions as verbal aggression, impoliteness (either mock or genuine), as well as teasing and banter (Culpeper 2011, Dynel 2008, Kádár, Parvaresh & Ning 2019). While offensiveness can facilitate rapport-building and mitigation of face-threatening acts in close or intimate relationships (Dynel 2021), in public discourse, the offence initiator has to deal with a greater distance from the target of their offence. In addition, the recipient of the offence is underspecified – in the age of social media it can change within minutes – which makes it nearly impossible to make provisions for an intended audience.

Utterances can be taken as humorous or not humorous, either because you don't get the reason for the potential humour or because you get it but you don't find it funny in the particular situation (Hale 2018). Likewise, particular

utterances can be taken as offensive by some people and not offensive by others, either because they don't know it is (potentially) offensive (i.e. they are unaware that it could be taken as a social norm violation by some group regardless of whether they belong to that group), or because it does not cross some threshold for offensiveness for them in the specific situation. For example, words which would be offensive slurs if used by an out-group member can be adopted and taken as terms of affection within a group (Burnett 2020). Following Haugh (2015) we take a recipient-centric perspective on offence, in a similar vein to Tsakona's (2013) approach to humour. As with humour, we consider whether an utterance was intended to be taken as offensive (by some person or group) or not to be orthogonal to the question of whether some person or group actually did take offence at it, with offence taking itself a pragmatic action (McTernan 2023), and thus necessitating explanation by any comprehensive theory of pragmatics.

Much previous research on offensive humour (Haugh 2015, Culpeper 2011) has focussed on the social norm violations which must be present for something to be taken as offensive, regardless of whether the offensive content is couched or intended to be taken as humorous. Indeed, whether you get why a joke is supposed to be funny is at least potentially independent of whether you think it is offensive. Despite this, although offence hasn't been studied as extensively as humour (Haugh 2015), they appear to arise in similar contexts. For instance, in a Meaney et al. (2021) shared computational task, humorousness ratings of jokes are correlated with ratings of personal and general offence. Our examples, however, suggest that an individual does not usually simultaneously find something humorous and take personal offence. Rather, understanding an utterance as an instance of offensive humour seems to rely at least in part on an awareness of the potential for offensiveness but often from the projected standpoint of a group to which the listener does not identify. Such fine-grained considerations are beyond the scope of this paper and we leave them aside for now.

Messages that may be perceived as controversial are usually dispreferred (Robinson 2004), and hence avoided in public discourse – think of the politician avoiding to answer an uncomfortable question. However, sometimes, when a sender seeks to gain the support of a group to which generally controversial opinions or values are appealing, it can be beneficial to send a controversial message. In such cases we often encounter the dogwhistle communication mentioned in the previous section, where an in-group perceive a controversial message while the rest of the audience perceive another, non-controversial message. In dogwhistle discourse the mismatch in topoi is perhaps more apparent than elsewhere – for those in the know (Sayeed et al. 2024).

While it is true that different topoi can be evoked in the minds of different groups of hearers of the same message for nefarious ends, as is (presumed to be) the case for dogwhistles, these generally available mechanisms for interaction can also be used for humour. Maraev et al. (2021) cast jokes in terms of topoi (as a granular resource to account for different ways of opposing joke scripts) and enthymemes (as incomplete arguments occurring in a dialogue or text, and involving one or more topoi; Breitholtz 2020) that arise from specific interac-

tional experiences. In this analysis, humour arises because there are two or more conflicting topoi which could underpin the enthymematic utterance (see also Breitholtz & Maraev 2019). We believe that the same type of analysis can be extended to questions of offensiveness, with the difference being that the topoi in (potentially) offensive contexts which clash in some way are necessarily about social norms.

In multiparty interaction, it is not uncommon for one to be aware that certain members of one's audience may not pick up on all the inferences that are available to others in the audience (as in the case of in-jokes, for example). This is often exploited for less politically troubling aims than dogwhistles in the case of children's films, which often include jokes for the adults accompanying their children to the cinema. These rely on inferences based on topoi that are (it is hoped) inaccessible to the children, often relying on innuendo, such as when Anna says to Christoff in Frozen "Foot size? Foot size doesn't matter".

Whilst such jokes and dogwhistles intentionally exploit the availability of more than one potential topos to underpin any enthymematic utterance, such effects can also arise unintentionally. This is especially the case in our current hyper-connected world, where our assumptions about who our listeners are, and thus the topoi they will draw on when interpreting an utterance do not necessarily hold. An extreme example of this was the well publicised case of Justine Sacco. As discussed in Millas (2016), in 2013 Justine Sacco tweeted an ill-advised joke to her 170 followers "Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!". In the ensuing twitter storm, Justine Sacco was branded a racist and was fired from her job.[5]

Once produced, the joke is left to the interpretation of the listeners – who all bring different preferences, experience, cultural expectations et cetera to the table. Even when we try to take into account other people's perspectives, we all interpret things through the prism of our own experience and expectations (which also includes how much we try to understand how different other people's perspectives are). In other words – the utterances were designed to be interpreted according to one topoplex but were put in the context of another topoplex that is different in some ways that are crucial to the interpretation. In the case of taking offence, as in the Justine Sacco example, we can see that speakers can be held responsible for the different interpretations other unexpected groups of listeners may make – even if they did not themselves intend such an interpretation (Haugh 2013, Leth 2019).

3   Topoi based analysis

In this section, we introduce our examples and analyse them in terms of the topoi that different audiences may access, leading to different interpretations of both offensiveness and humour. Following Maraev et al. (2021), our analysis is

---

[5]See also e.g. Vladimirou, House & Kádár (2021) for how social media can escalate aggression.

intended to fit into the theoretical dialogical account of Ginzburg (2012) and Cooper (2023).

3.1   We could shoot him

Our first example marks potentially offensive content as a joke by means of making the content hard to believe. It comes from David Attenborough's 2016 interview with Radio Times journalist Emily Maitlis to promote his new nature documentary series (the Radio Times is a British TV-listings magazine).[6] According to the article, when asked about what he thinks about soon having "a man in charge of America who believes climate change is a Chinese hoax" (in reference to previous comments from the then presidential candidate Donald Trump), Attenborough's reply is as follows:

> Attenborough's head is in his hands but his response is curiously phlegmatic. Or perhaps pragmatic. "Yes, I know. Well, we lived through that with earlier presidents – they've been equally guilty... But what alternative do we have? Do we have any control or influence over the American elections? Of course we don't. [sotto voce] We could shoot him," he jokes. "It's not a bad idea..." He catches my eye and giggles.

In this case, the primary audience are British people for whom David Attenborough is a well-respected naturalist, often referred to as a "national treasure" or "everyone's favourite Grandfather". Additionally, the journalist puts in a lot of effort to indicate that the comment was both made as and taken as a joke in the immediate context ("[sotto voce]"; "he jokes"; "He catches my eye and giggles").

The initial responses to the interview were from this expected audience and clearly report what David Attenborough says as a joke to emphasise his dislike of then presidential candidate Donald Trump, as the below quotes from online news site *Mashable* and British mainstream newspaper *The Independent* demonstrate.

> It's important to stress that Attenborough was obviously just joking with that last bit ["We could shoot him... It's not a bad idea..."] – the interviewer writes "he catches my eye and giggles" after the quote – but one thing's clear: Attenborough is obviously not keen on the idea of Trump being in charge of things.[7]

> Adopting a hushed tone, he said: "We could shoot him, it's not a bad idea" before chuckling.[8]

---

[6] https://www.radiotimes.com/tv/documentaries/david-attenborough-on-planet-earth-ii-brexit-and-the-future-of-humanity/

[7] https://mashable.com/article/david-attenborough-donald-trump-shoot-him

[8] https://www.independent.co.uk/news/people/david-attenborough-donald-trump-shoot-radio-times-interview-michael-gove-a7390476.html

In sum, to an audience that views the likelihood of Attenborough actually shooting anyone at all as extremely low, if not zero, the suggestion of shooting Trump is interpreted as obvious sarcasm, drawing on the topos discussed in Section 2.1 that if x says p and p is clearly not the case, then x is not being serious. However, the expected audience, who might have more similar views on Trump's potential presidency, and a background cultural context where shooting people is not a daily occurrence, were not the only ones to react to the article.

American Trump supporters picked up on Attenborough's comments and interpreted them from their own context. They may lack topoi about David Attenborough being well-known and a national treasure and have a different set of topoi regarding shooting people. Such reactions were unsurprisingly reported differently in English and American contexts:

> Rather than revelling in their victory, Trump supporters are threatening a British national treasure… Supporters of America's new President-elect sent the beloved broadcaster violent threats, including: 'How about we put 2 in the back your head'.[9]

> Whether the U.S. Secret Service investigates this very public threat against Trump is unlikely as Attenborough lives in the U.K. and is not likely to carry out any shooting. But suggesting violence against a presidential candidate is a crime so he may find himself in hot water with his sponsors like the BBC and National Geographic…

> Attenborough is probably best known for being confused with his brother Richard Attenborough.[10]

3.1.1 Mismatching topoi

For the David Attenborough example, the two different interpretations from the UK versus the US can be analysed in terms of the readily available topoi available to each group.[11] There are (at least) two dimensions that are relevant for the interpretation – how realistic the scenario of a shooting is (based on your culturally conditioned world knowledge) and to what degree you think Attenborough is serious in his suggestion (based on your knowledge of and attitude towards his persona).

The enthymeme uttered by Attenborough can be represented as (12).

$$\frac{\text{How can we stop Trump}}{\text{We could shoot him}} \tag{12}$$

---

[9] https://metro.co.uk/2016/11/17/sir-david-attenborough-sent-death-threats-after-saying-we-could-shoot-donald-trump-6263850/

[10] https://us.blastingnews.com/news/2016/11/sir-david-attenborough-proposes-shooting-donald-trump-to-save-climate-001224351.html

[11] Note that we gloss these as the typical UK versus US populations, but of course this is a simplification for expository purposes and individual variation will not necessarily fall only along this geographical dimension.

For both the UK and the US audience, we assume that there is a readily available general topos which can underpin such an enthymeme, sketched in (13), which raises additional salient common topoi including things like (14), paraphrasable as (the obviously defeasible) 'if you shoot someone, you kill them'. This chain of reasoning led some commentators to report Attenborough as inciting murder.

$$\frac{\text{stop person x}}{\text{shoot person x}} \qquad (13)$$

$$\frac{\text{shoot person x}}{\text{kill person x}} \qquad (14)$$

While such topoi are available to both a UK and US audience, they are presumed to be less salient and generally applicable in the UK context. This is because of differences in the cultural contexts in the UK and the USA. In the USA many people own guns, and shootings (which often lead to death) are common, whilst in the UK gun ownership is largely restricted and fatal shootings are much rarer.[12] This means that the topos in (14) is likely to be more easily accessible to those in the USA, and also have a higher probability of being reliable, in the absence of other contextual cues to the contrary.

In the British case, the topoplex associated with Attenborough's persona includes topoi such as Attenborough is nice (15) (as evidenced by his 'National Treasure' status, and modulated by the audience's attitude towards nature, and his position as a respected environmentalist) and nice people don't kill people (16).

$$\frac{\text{x is David Attenborough}}{\text{x is nice}} \qquad (15)$$

$$\frac{\text{x is nice}}{\text{x would not kill someone}} \qquad (16)$$

This being the case, some extra inferential work must be done in order to make sense of the utterance on the fly, since we otherwise have a conflict in the conclusions, such that David Attenborough has apparently said we should shoot someone which leads to the conclusion that we should kill them, whilst simultaneously having a persona that leads to a conclusion that he would not kill them. The topoi about David Attenborough's persona, which apparently undermine the seriousness of his comment are both more available to some listeners than others and more or less acceptable.

Further, topoi that a group entertains about the *target* (group or individual) of the potentially offensive utterance can also affect one's willingness to find it humorous or offensive. In this case, Trump supporters would have positive topoi associated with Trump (analogous to (15)), thus making a conclusion of someone wanting to shoot him even less palatable. We do not spell these topoi out here.

---

[12]In the most recently available data there are 4.054 deaths by shooting per 100000 people in the USA compared to 0.047 in the UK, i.e. they are 100 times more likely in the USA.

Even if conflicting topoi are accessible, certain members of the audience may choose to reject them, taking offence by apparently taking the utterance at face value despite being aware that it was intended as a non-serious utterance. This means that it is not just the availability of the conflicting topoi, but also the weight given to them which determines both whether or not you understand the joke **and** whether or not you find it funny and/or offensive.

To address this more formally, as discussed in Section 2.2, we can draw on Noble, Breitholtz & Cooper's (2020) approach to treat topoi as a set of probability distributions associated with a certain persona. In the David Attenborough case, the UK audience is likely to project "environmentalist" or "humanist" personae for which the topos of an acceptable way to stop someone by means of shooting them (13) is very unlikely to be evoked. In other words, it will be incongruous with what we are likely to hear from them. Or, to be more precise, it is very unlikely to underpin the enthymeme (12) at face value, therefore it needs extra work to be accommodated, specifically, it should be accommodated as a non-serious claim.[13]

3.2   Alla mina kamrater

Our second example comes from a series of offences performed by the Swedish comedian Kristoffer Svensson (Kringlan) while being interviewed for a chatty Swedish podcast called "Alla mina kamrater" (All my comrades) (Franzén, Jonsson & Sjöblom 2021). During the podcast episode, Kringlan uttered a number of violent, sexual, misogynist threats towards Åsa Linderborg, the editor of the daily Swedish tabloid newspaper Aftonbladet's culture pages. Kringlan felt that a review of his book published by the paper (notably not written by Linderborg) was unnecessarily critical. Kringlan's "norm-breaching" angry rant, a part of which is shown in (17), was taken by the podcasters as being excessive, but nevertheless as a joke. In laughing at Kringlan's offensive comments the podcast hosts were also implicated in the offensiveness of the discussion.

(17)   From Franzén, Jonsson & Sjöblom (2021)
    182   Kringlan: Men ja e liksom bortom Twitterbråk ja ska liksom dra
                   'But I'm like beyond Twitter fights like I'm gonna put'
    183   Kringlan: en yxa i fittan på Åsa [Linderborg]
                   'an axe in the cunt of Åsa Linderborg'
    184   Fritte: [WOAH!]
                   'WOAH!'
    185   Several: [He]hehe
                   'Hehehe'

In this case again, there is a mismatch between the immediate audience (the podcast hosts) who were all young, white, Swedish males and the wider

---

[13]Note that the incongruity which arises in such cases for the intended audience doesn't keep the message of the joke from being understood. The joke is used to make Attenborough's point about being an environmentalist and anti-Trump. Such a "tendenziös" (Freud 1905) agenda can be inferred from other topoi invoked by the joke which are beyond the scope of this article.

audience (the podcast listeners). As this is a prerecorded format, the podcast hosts (and editorial team) actually had to decide that this should be released. Perhaps surprisingly, they did not anticipate that the wider audience (including many women for whom the misogynistic nature of the language was taken to be offensive) would not consider Kringlan's comments to be a joke, and/or assign shared responsibility for the comments to the podcast hosts. Following a media outcry, both Alla Mina Kamrater and Kringlan were cancelled. One week after the initial broadcast, the Alla Mina Kamrater team announced they would not be recording any more podcasts, and Kringlan's publisher, Aftonbladet and Sweden radio also cancelled his contracts, including deleting his contributions to a review of the year show.

Figure 1: The "Alla Mina Kamrater" podcast hosts

Figure 2: Kristoffer Svensson (Kringlan)

In the Kringlan case, we see a superficially similar structure to the David Attenborough one. However, unlike (13) there is no readily available common topos from wanting to punish someone to putting an axe in their vagina. Furthermore, there are many readily available negative topoi around explicit sexual violence, and for most of the audience, there is no preexisting notion of Kringlan's persona that can make these unexpected or allow them to draw incompatible conclusions which may prompt accommodation. In contrast the statement itself may raise additional topoi such as (18), which make an offence-taking reaction more likely, assuming you also have common social norm topoi available that if a person is a misogynist that person is not nice.[14]

$$\frac{\text{x made a misogynist statement}}{\text{x is a misogynist}} \tag{18}$$

In this case, the podcast hosts may have had some additional topoi available based on Kringlan's persona and the disproportionality of his violent misogyny that provided the necessary context to make the utterance appear amusing.

[14]Note that while this might be a commonly held topos in the West, the existence of the manosphere and incels for example, mean it is not universally accepted (Farrell et al. 2019).

Both the David Attenborough and the Kringlan cases present an exaggerated unrealistic, offensive action that the respective audiences are expected to understand are not intended to be carried out. It is precisely this mismatch between what is said and what is realistically going to happen that is supposed to be the source of the humour. However, to find the offensive exaggeration funny requires additional knowledge about why it is unlikely that the particular speaker is being serious, which works in the Attenborough case – for a British audience – because of the preexisting topoi associated with him. The persona associated with Kringlan (and the Alla Mina Kamrater podast hosts) on the other hand, does not preclude his actually being a misogynist and really wanting to cause harm of a sexual nature to a woman who he believes has slighted him.

As can be seen from these examples, offensiveness can not be attributed to speakers independently of the audience to the offensive utterance. It is the expectations of the hearers (Tayebi 2016) – which we model as the topoi they have available in the context – which dictates whether offence is taken.

3.3   Helg seger!

In September 2022, at the Swedish general election, the Sweden Democrats (SD) party did particularly well, receiving 20.5% of the vote, making them the second largest party. Although SD deny that they are a "far-right" political party, it is well-known that many of the party's founders (around 9 out of 30 according to Larsson & Ekman 1999) had direct associations to known neo-Nazi fascist organisations. During the watch party organised by SD, party member and working politician Rebecka Fallenkvist was asked by a reporter about her thoughts on the positive emerging results.

As she replied, Fallenkvist raised her left arm (as shown in Figure 3) and used the non-idiomatic phrase "Helg Seger", which translates as "weekend victory". However, this phrase is also phonetically very similar to "Hell Seger", the Swedish version of the German Nazi salute "Sieg Heil". This would have been well-known to those with neo-Nazi sympathies and was the focus of extensive negative media coverage following the event. When questioned directly about her use of the phrase "Helg Seger!" Fallenkvist first claimed that this was a joke aimed at easily provoked leftist media. However, she later claimed that any similarity with the Nazi salute was accidental and unintentional. Indeed, immediately after producing the phrase "Helg Seger!", Fallenkvist reformulated it to the more plausible "Segerhelg" (victory weekend).

Fallenkvist's utterance is an example of how offensive things can be hinted at in a "humorous" (and potentially deniable) way – as a slip of the tongue. As with dogwhistles, using humour to evoke different (and conflicting topoi) in the minds of different audience members can be used to communicate an offensive message to a subset of one's audience.

This example highlights the relevance of cultural context to humour, and also the fact that different degrees of offensiveness can be associated to the same utterance in different groups. When Fallenkvist utters "Helg seger" she

Figure 3: Rebecka Fallenkvist being interviewed about the emerging election results

is at a party surrounded by people to whom this particular Nazi reference is well known, and potentially offensive, but not to the degree that is outside far right circles. In some respects this example is more similar to a dogwhistle, and maybe that was what was intended, as the topos that the phrase raises (19) is potentially not well known to many people outside far right circles (particularly non-Swedes).

$$\frac{\text{Hell Seger = Sieg Heil} \quad \text{x says Helg Seger}}{\text{x is using a Nazi Salute}} \quad (19)$$

However, Fallenkvist also raises her arm in a gesture that appears similar to the well known Nazi salute, which makes salient a topos that is available to a wider audience.

$$\frac{\text{x raises their arm in a particular way}}{\text{x is using a Nazi Salute}} \quad (20)$$

Both of these topoi lead to reasoning such as that depicted in (21), which is normatively considered bad.

$$\frac{\text{x uses a Nazi salute}}{\text{x is a Nazi}} \quad (21)$$

However, it was clearly known to enough people to be heavily discussed in the media. Further, we hypothesise that the Nazi salute is to some extent taboo

in far right circles as well as in society at large (particularly for appearance' sake). The persona presented by Fallenkvist to those outside of SD circles also contains far right connotations, by association to her representing SD.

$$\frac{\text{x is Rebecka Fallenkvist}}{\text{x is far-right}} \tag{22}$$

$$\frac{\text{x is far-right}}{\text{x is a Nazi}} \tag{23}$$

It is well-known that (23) often applies, at least according to progressive left rhetoric. It is particularly interesting therefore to consider how Fallenkvist reacted to criticism from the media. Although she later claimed that it was accidental and unintentional (thus attempting to leverage some of the deniability which is a characteristic feature of dogwhistles for examples), her initial response was that it was a joke aimed at the left. This, of course, makes her later claims that it was just a slip of the tongue less plausible. Note that this does not mean that she did actually act intentionally as a joke towards the easily provoked left. It is possible that in the context where she was initially challenged she may have thought that the audience would sympathetic to a joke at the left's expense, and that it was therefore potentially advantageous to claim post-hoc responsibility to this intention (Gregoromichelaki et al. 2011), even if it had actually been a genuine slip of the tongue at the time. This interpretation was not widely applied by commentators.

4 Future steps

Based on the analysis of these three cases, we can sketch out a procedure for analysis of similar cases in future work. Note that for the purposes of this analysis schema, we remain agnostic as to individual intentions behind a response, which is of course an important additional dimension, but outside the scope of this paper. What matters here is rather whether we can map the behaviours of different hearers to different sets of topoi. In practice of course, people may actually share the same topoi as another hearer but choose to be offended (and thus foreground a different topoplex) as a political strategy.

1. Identify instance of potentially offensive humour
2. Identify topoi salient to a particular hearer interpreting the instance based on

   - (perceived) persona of speaker
   - characteristics of target group or person
   - (perceived) relationship between speaker and target
   - identity of hearer
   - relationship between hearer and target

3. Repeat for a different hearer to identify mismatches. Is the likelihood of taking offence affected by some hearers having

17

- additional topoi?
- no access to particular topoi?

This fine-grained level of topoi for explaining humour and offence makes it possible to evaluate our claims empirically. A next step once different examples are identified is an experimental intervention which uses systematically different formulations of the offence/humour and with different projected personae (e.g., imagining it being produced by different speakers).

We are planning to conduct experimental studies where we will attempt to "break" jokes by altering the topos evoked or the time when it is evoked. We would also expect potentially offensive utterances to be more or less likely to cause offence depending on the apparent persona of the speaker, as well as e.g. the level of explicitness. For example, it might be taken as more offensive if someone who inhabits a persona of which it is an essential trait to be in opposition to Trump, who suggested the shooting, for example Kamala Harris (Trump's opponent in the 2024 US presidential election).

In a similar way, a more graphic form of suggesting the shooting by Attenborough could carry a higher degree of offensiveness (e.g., "[sotto voce] We could blow his brains out…").

We also plan to extend previous formal models of humour (Maraev et al. 2021, Breitholtz & Maraev 2019) to include offensiveness, but also gradience with regard to both humour and offensiveness, reflecting the results of the preceding experiments.

5 Conclusions

Controversial and potentially offensive messages can be communicated in a number of ways by exploiting the topoi which are evoked in the mind of the hearers. Any analysis of this must take into account that different audiences have different "minds" with different common assumptions, cultural experience and social relationships, etc. We propose that these sets of assumptions can be modelled as sets of topoi, which we call topoplexes. We consider these to be the basis of different interpretations of utterances causing them to be interpreted as humorous and/or offensive. In this paper we looked at three specific instances of offensive humour. We emphasise that humorous utterances may be used to convey a potentially controversial message in an indirect way that may relieve the speaker of accountability. We also point out a link between potentially offensive humour and (political) dogwhistles.

We believe that this type of fine-grained analysis (Noble, Maraev & Breitholtz 2022, Breitholtz & Maraev 2019) can pave the way for precise and explainable computational models of offensive humour (in contrast to the "black box" reasoning of current large language models). Overall, we believe that explaining offensive humour using general principles of dialogue provides a more precise interactive account which can account for the subjectivity of both offence and humour, and mismatches between its interpretations and appraisals

by different listeners, as well as larger social communities. As discussed in Aly & Simpson (2019) this type of mismatch can lead one group to potentially claim that someone taking offence at something allegedly said in jest is an example of political correctness gone mad. This raises a new topoplex and serves to project a negative "killjoy" persona on anyone who (potentially legitimately) took offence, illustrating the complex interaction between humour and offence.

This paves a way towards a general theory of human reasoning in dialogue which can encompass both everyday examples (such as Arthur and Evelyn's discussion of the salt on the roads), with examples from public discourse where the generally available interactive reasoning mechanisms are exploited for humorous effect or to conceal a potentially controversial message – or, as is the case in our examples, potentially both at the same time.

References

Aly, Waleed & Robert Mark Simpson. 2019. Political correctness gone viral. In *Media ethics, free speech, and the requirements of democracy*, 125–143. New York: Routledge.

Anscombre, Jean-Claude. 1995. La théorie des topoi: Sémantique ou rhétorique? *Hermés* 15.

Aristotle. 2007. *On rhetoric, a theory of civic discourse* (translated by George A. Kennedy). (Original work published ca. 340 B.C.E.) Oxford: Oxford University Press.

Ball, Linden J & Valerie A Thompson. 2017. *International handbook of thinking and reasoning.* New York: Routledge.

Beaver, David & Henk Zeevat. 2007. Accommodation. *The Oxford handbook of linguistic interfaces.* 503–536.

Bell, Allan. 1984. Language style as audience design. *Language in Society* 13(2). 145–204.

Breitholtz, Ellen. 2020. *Enthymemes and topoi in dialogue: The use of common sense reasoning in conversation.* Leiden, The Netherlands: Brill.

Breitholtz, Ellen & Christine Howes. 2020. Communicable reasons: How children learn topoi through dialogue. *Proceedings of the 24th Workshop on the Semantics and Pragmatics of Dialogue.*

Breitholtz, Ellen, Christine Howes & Robin Cooper. 2023. All the more reasons: Mismatches in topoi in dialogue. *Journal of Pragmatics* 217. 172–184. `https://doi.org/https://doi.org/10.1016/j.pragma.2023.09.006`.

Breitholtz, Ellen & Vladislav Maraev. 2019. How to put an elephant in the title: Modelling humorous incongruity with topoi. *Proceedings of the 23rd Workshop on the Semantics and Pragmatics of Dialogue.*

Burnett, Heather. 2020. A persona-based semantics for slurs. *Grazer Philosophische Studien* 97(1). 31–62.

Cooper, Robin. 2023. *From perception to communication: A theory of types for action and meaning.* Oxford: Oxford University Press.

Culpeper, Jonathan. 2011. *Impoliteness: Using language to cause offence.* Vol. 28. Cambridge: Cambridge University Press.

Davies, Christie. 1990. *Ethnic humor around the world: A comparative analysis.* Bloomington: Indiana University Press.

Davies, Christie. 1998. *Jokes and their relations to society.* Berlin, Boston: De Gruyter Mouton. `https://doi.org/doi:10.1515/9783110806144`.

Ducrot, Oswald. 1980. *Les échelles argumentatives.* Paris: Les Éditions de Minuit.

Ducrot, Oswald. 1988. Topoï et formes topique. *Bulletin d'Études de la Linguistique Française* 22. 1–14.

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

Dynel, Marta. 2021. COVID-19 memes going viral: On the multiple multimodal voices behind face masks. *Discourse & Society* 32(2). 175–195. `https://doi.org/10.1177/0957926520970385`.

Farrell, Tracie, Miriam Fernandez, Jakub Novotny & Harith Alani. 2019. Exploring misogyny across the manosphere in reddit. *Proceedings of the 10th ACM Conference on Web Science.* 87–96.

Franzén, Anna G., Rickard Jonsson & Björn Sjöblom. 2021. Fear, anger, and desire: Affect and the interactional intricacies of rape humor on a live podcast. *Language in Society* 50(5). 763–786. `https://doi.org/10.1017/S0047404520000615`.

Freud, Sigmund. 1905. *Jokes and their relation to the unconscious.* London: Hogarth.

Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation.* Oxford: Oxford University Press.

Gregoromichelaki, Eleni, Ruth Kempson, Matthew Purver, Greg J. Mills, Ronnie Cann, Wilfried Meyer-Viol & Patrick G. T. Healey. 2011. Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse* 2(1). 199–233.

Hale, Adrian. 2018. "I get it, but it's just not funny": Why humour fails, after all is said and done. *The European Journal of Humour Research* 6(1). 36–61.

Haugh, Michael. 2013. Speaker meaning and accountability in interaction. *Journal of Pragmatics* 48(1). 41–56. `https://doi.org/https://doi.org/10.1016/j.pragma.2012.11.009`.

Haugh, Michael. 2015. Impoliteness and taking offence in initial interactions. *Journal of Pragmatics* 86. Interdisciplinary perspectives on pragmatics: A Festschrift for Jonathan Culpeper, 36–42. `https://doi.org/https://doi.org/10.1016/j.pragma.2015.05.018`.

Henderson, R. & Elin McCready. 2018. How dogwhistles work. In Sachiyo Arai, Kazuhiro Kojima, Koji Mineshima, Daisuke Bekki, Ken Satoh & Yuiko Ohta (eds.), *New frontiers in artificial intelligence*, 231–240. Cham: Springer International Publishing.

Jackson, Sally & Scott Jacobs. 1980. Structure of conversational argument: Pragmatic bases for the enthymeme. *Quarterly Journal of Speech* 66(3). 251–265.

Kádár, Dániel Z, Vahid Parvaresh & Puyu Ning. 2019. Morality, moral order, and language conflict and aggression: A position paper. *Journal of Language Aggression and Conflict* 7(1). 6–31.

Kuipers, Giselinde. 2015. *Good humor, bad taste: A sociology of the joke.* Berlin: Walter de Gruyter GmbH.

Larsson, Stieg & Mikael Ekman. 1999. *Sverigedemokraterna - den nationella rörelsen.* Ordfront förlag.

Leth, Palle. 2019. Is there any use for a notion of the correct interpretation of an utterance? *Philosophical Insights into Pragmatics* 79. 83.

Lev-Ari, Shiri. 2018. The influence of social network size on speech perception. *Quarterly Journal of Experimental Psychology* 71(10). 2249–2260.

Lewis, David. 1979. Scorekeeping in a Language Game. *Journal of Philosophical Logic* 8. 339–359.

Linell, Per. 2009. *Rethinking language, mind, and world dialogically.* Charlotte, NC: IAP.

Maraev, Vladislav, Ellen Breitholtz, Christine Howes, Staffan Larsson & Robin Cooper. 2021. Something old, something new, something borrowed, something taboo: Interaction and creativity in humour. *Frontiers in Psychology.* https://www.frontiersin.org/articles/10.3389/fpsyg.2021.654615/full.

Mazzocconi, Chiara, Vladislav Maraev & Jonathan Ginzburg. 2018. Laughter repair. *Proceedings of the 22nd Workshop on the Semantics and Pragmatics of Dialogue.* 16–25.

McTernan, Emily. 2023. *On taking offence.* Oxford: Oxford University Press.

Meaney, J. A., Steven Wilson, Luis Chiruzzo, Adam Lopez & Walid Magdy. 2021. SemEval 2021 task 7: HaHackathon, detecting and rating humor and offense. *Proceedings of the 15th International Workshop on Semantic Evaluation.* https://doi.org/10.18653/v1/2021.semeval-1.9.

Millas, Andrea Pérez. 2016. Forget whipping: Log in to Twitter for ignominy. *Txt.* Retrieved from https://hdl.handle.net/1887/42725, 128–131.

Milroy, James & Lesley Milroy. 1978. Belfast: Change and variation in an urban vernacular. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 19–36. Baltimore: University Park Press.

Musolff, Andreas. 2004. *Metaphor and political discourse: Analogical reasoning in debates about europe.* Basingstoke, UK: Palgrave MacMillan.

Noble, Bill, Ellen Breitholtz & Robin Cooper. 2020. Personae under uncertainty: The case of topoi. In *Proceedings of the probability and meaning conference (PaM)*, 8–16.

Noble, Bill, Vladislav Maraev & Ellen Breitholtz. 2022. Probabilistic pragmatics: A dialogical perspective. In J.-P. Bernardy, R. Blanck, S. Chatzikyriakidis, S. Lappin & A. Maskharashvili (eds.), *Probabilistic approaches to linguistic theory*, 227–258. Stanford, California, USA: CSLI Publications.

Robinson, Jeffrey D. 2004. The sequential organization of "explicit" apologies in naturally occurring English. *Research on Language and Social Interaction* 37(3). 291–330.

Rosengren, Mats. 2002. *Doxologi: En essä om kunskap.* Copenhagen: Retorik-förlaget.

Sayeed, Asad, Ellen Breitholtz, Robin Cooper, Elina Lindgren, Gregor Rettenegger & Björn Rönnerstrand. 2024. The utility of (political) dogwhistles – a life cycle perspective. *Journal of Language and Politics.*

Stalnaker, Robert. 2014. *Context.* Oxford: Oxford University Press.

Tayebi, Tahmineh. 2016. Why do people take offence? Exploring the underlying expectations. *Journal of Pragmatics* 101. 1–17.

Tsakona, Villy. 2013. Okras and the metapragmatic stereotypes of humour: Towards an expansion of the GTVH. In Marta Dynel (ed.), *Developments in linguistic humour theory*, 25–48. Amsterdam: John Benjamins.

Van Dijk, Teun A. 2008. Context theory and the foundation of pragmatics. *Studies in Pragmatics* 10. 1–13.

Vladimirou, Dimitra, Juliane House & Dániel Z Kádár. 2021. Aggressive complaining on social media: The case of #muckymerton. *Journal of Pragmatics* 177. 51–64.

Warren, Caleb & A Peter McGraw. 2016. Differentiating what is humorous from what is not. *Journal of Personality and Social Psychology* 110(3). 407.