

What Is It to Be Responsible for What You Say?

EMMA BORG

Abstract

In asserting something I incur certain kinds of liabilities, including a responsibility for the truth of the content I express. If I say ‘After leaving the EU, the UK will take back control of c. £350 million per week’, or I tell you that ‘The number 14 bus stops at the British Museum’, I become liable for the truth of these claims. As my audience, you could hold me unreliable or devious if it turns out that what I said is false. Yet this socio-linguistic practice – of acquiring and ascribing ‘linguistic liability’ – is complicated, especially given philosophical distinctions between the various different kinds of contents people can express (am I liable, for instance, for the claim that the number 14 bus stops at the British Museum *today* or only *usually*?). This paper explores the different kinds of contents speakers might be taken to express, arguing that our practices around linguistic liability (including in legal disputes) reveal a crucial role for a notion of context-independent, literal meaning attaching to words and sentences. These practices thus vindicate what philosophers tend to term ‘minimal semantic content’.

1. Introduction

In 2017 US President Donald Trump stated that the crowd at his inauguration appeared to stretch ‘all the way back to the Washington Monument’. Later Trump’s Press Secretary, Sean Spicer, doubled down on this claim, telling reporters that ‘This was the largest audience ever to witness an inauguration – period’. Many press outlets, however, disputed these statements. An article in the UK newspaper *The Guardian* (22 January 2017), for instance, reproduced photographs from the event and concluded that ‘the evidence certainly seems to challenge Trump’s assertion’, while a *New York Times* piece the following day went further, stating that Mr Spicer’s claim was ‘clearly shown to be false’ by the photographs.

Setting aside the political dimensions of these exchanges, the events lay bare a fundamental aspect of our linguistic practices: hearers assume that speakers aim to assert the truth and that speakers can be held culpable if what they say turns out not to be true. This connection between assertion and truth seems fundamental to our linguistic practices: it is only because an audience can assume that

the speaker is aiming to assert what is true that hearers can use linguistic testimony to inform them about, and direct their interactions with, the world. If I want a cookie and you say ‘There are cookies in the jar’, I can act on what you say, using it to direct my cookie-hunting actions, only because I assume that you are telling the truth.

The interplay between meaning, communication, and truth is also fundamental to a number of philosophical theories. For instance, the Oxford philosopher of language Paul Grice held that, for communication to be possible at all, interlocutors need to be able to assume that their conversational partners are *cooperative*, i.e., that they are aiming to convey appropriate and true information (Grice, 1989). More recently another Oxford philosopher, Tim Williamson, has argued that truth provides the ‘norm of assertion’, whereby a speaker can only assert what they know to be true (Williamson, 1996). However, although the requirement to assert what is true seems philosophically and practically fundamental, on closer inspection the requirement turns out to be more complicated than we might initially have supposed. For in order to hold a speaker responsible for the truth of what they say there are two things we first need to know:

- (1) What did the speaker say (i.e., what content did they assert)?
- (2) What kind of responsibility is in play?

Answering these two questions, however, at least from the theoretical point of view, is far from straightforward. The aim of this paper is to explore this complexity and show how a proper understanding of our practices of holding speakers responsible for what they say can help to advance an entrenched philosophical dispute.

I start, in §2, by examining question (1), asking what kinds of content a speaker can be held to express in uttering what they do. As we will see, the three most plausible answers here line up with three different approaches in philosophy of language: first, so-called ‘semantic minimalism’ (which, as we will see, claims that well-formed declarative sentences express genuine, truth-evaluable content without significant appeal to the context in which they are uttered), second ‘contextualism’ (which holds that the key meanings are somewhat contextually enriched contents, ones which capture the content implicit in what a speaker says), and finally an approach drawn from the work of Paul Grice which focuses on what a speaker conversationally implies by what they say. One immediate question, however, concerns the relationship between these three different approaches: are they competitor accounts, or alternatively are they all needed for a complete account of what a speaker says? That is to say, do we really need to posit all three kinds of content which

What Is It to Be Responsible for What You Say?

the three different accounts posit, or is a more philosophically parsimonious approach possible? In this vein, advocates of contextualism have long argued that we can do without the kind of content proposed by semantic minimalists, that sentences stripped of contextual input do not (generally or perhaps ever) express a complete or worthwhile content. Furthermore (again, as we will see), minimalist content might be thought to rest on shaky foundations anyway since it requires the notion of context-insensitive word meaning, which itself looks pretty problematic. Trying to respond to this latter worry on behalf of the minimalist will push us, in §§3–4, to examine the notion of word meaning and the ways in which people arrive at an understanding of the content their words express. I'll suggest that words can be thought to have standing, context-insensitive meanings and that (given the right combination of subjective and objective approaches to linguistic meaning) those meanings are ones interlocutors can grasp. Thus, I'll argue that the existence of minimal *sentence-level* content should not be rejected on the basis of worries about minimal *word* meaning. In §5 I turn to question (2) and argue that (contra contextualist objections) minimalist content (the content delivered by putting standing word meanings together in grammatically acceptable ways) is in fact *required*, since it plays a crucial role in our practices of assigning and accepting linguistic responsibility. Getting clear on the kinds of linguistic responsibility in play (answering question (2)) turns out, I argue, to vindicate the minimalist's claim concerning the need for minimal content.

2. What Does a Speaker Say?

Consider the following toy exchange:

Abbie: Do you want to have lunch?

Bishma: I've eaten.

To decide what content Bishma is responsible for expressing here, we need to know what he has said. Yet it seems different answers to this question are possible. A first option would be to focus on the words Bishma has uttered and take him to be expressing the content that results simply from putting words with just these meanings together in just this order. So, it seems that Bishma might be taken to have asserted simply the claim that *I, Bishma, have eaten*. Notice that this minimal claim is true just in case Bishma has eaten *something* at *some time* in the past, it doesn't tell us anything further about what

or when he's eaten (to see this, think about an utterance of 'I've eaten camel', where it's clearer that the speaker need not be committing to eating within a particular timeframe, see Sperber & Wilson, 1986, 189–90, or 'I've just eaten' which specifies a timeframe apparently without committing to a kind of thing eaten). If we understand Bishma as making this minimal assertion, it will turn out that he is a reliable and responsible speaker so long as he has eaten something at some point in the past, even if he hasn't eaten anything recently (i.e., in the run up to lunch today).

Alternatively, however, we might look to the wider context in which Bishma says what he does and treat some of those contextual features as affecting what Bishma says. For instance, given the context of Abbie's query about lunch, we might hold that Bishma asserts a richer content, something along the lines of *I have eaten lunch today* or *I have eaten recently*. Philosophers often label this kind of contextually enriched content an *explicature*, since it apparently makes explicit content that is implicit in the original utterance. In the above context, Bishma certainly doesn't convey the unrestricted claim that he has merely eaten at some time in the past, rather 'eaten' seems to convey the more contextually relevant idea that the speaker has *eaten recently*.

Third and finally, it seems we might look to an even more context-sensitive kind of content, focusing on the kind of move Bishma intends to make in the conversation. Assuming (as Grice suggested) that we view Bishma as aiming to be a cooperative conversational partner, it seems that we are entitled to assume that he is trying to provide an answer to Abbie's question. However, neither the minimal content (*Bishma has eaten*) nor the explicature content (*Bishma has eaten recently*) directly provides such an answer. We can, however, use the latter content to infer an answer: if Bishma has eaten recently it is reasonable to infer that *Bishma doesn't want lunch*. Thus we could hold that what Bishma says is a (polite) declining of the offer of lunch. Following Grice, philosophers tend to label this third kind of content an *implicature*, for it captures content that is merely implied by Bishma's utterance in this context. Implicature content is (usually) extremely context dependent: there are very many contexts in which an utterance of 'I've eaten' would not express a rejection of a lunch offer, but at least in the above context saying 'I've eaten' can be heard as saying *I won't have lunch with you*.¹

¹ A possible exception to the greater context-sensitivity of implicatures is Grice's category of Conventional Implicatures where words themselves

What Is It to Be Responsible for What You Say?

Depending on how much we allow considerations drawn from the context of utterance to influence our judgements, then, it seems that there are three distinct candidates available to play the role of what is said:

Minimal contextual input	Medium contextual input (explicature)	Maximum contextual input (implicature)
Bishma has eaten.	Bishma has eaten <i>recently</i> .	Bishma doesn't want lunch.

These three options line up with the contents focused on by three different philosophical schools of thought. The idea that there is an important, truth-evaluable level of content which can be recovered simply by looking to the words a speaker produces and the way they are put together (plus some very limited contextual input, such as determining the referent of 'I' in our toy example) is argued for by a position known as 'semantic minimalism' (see, e.g., Borg 2004, 2012). According to semantic minimalism, looking just to word meaning and grammatical structure yields a genuine, truth-evaluable content (i.e., a content which makes a claim which, when held up against the world, can be assessed for truth or falsity) and it is a level of content to which speakers and hearers are sensitive. However, advocates of minimalism also acknowledge that this kind of content is not usually the content which speakers express, or are taken to express, in communicative exchanges. Instead, minimalists allow that, during the cut and thrust of communication, the key content is more often some sort of pragmatically enriched content (i.e., explicature or implicature content).

Given this concession, however, so-called 'contextualists' (or 'pragmaticists') – those who focus on explicature content – have argued that there is no need to posit minimal content (see, e.g., Levinson, 2000, p. 231; Carston, 2008, p. 366). Contextualists suggest that word meaning plus structure (sometimes or perhaps always) yields only a partial, indeterminate meaning, one which hearers are licensed to

are held to convey a further implicated content. For instance, Grice held that 'but' literally means the same as 'and', with both expressing a simple conjunction, yet every utterance of 'but' carries a conventional implicature to the effect that there is some contrast between the two conjuncts (as in 'She was poor but honest'), see Grice (1989). However, the category of Conventional Implicatures has proved controversial (see Bach, 2006).

enrich (or modulate in other ways) in order to arrive at a content which is better suited to the communicative exchange which is taking place. A very influential approach along these lines is ‘Relevance Theory’ (proposed by Sperber and Wilson, 1986; Carston, 2002) which holds that the content encoded in the words speakers’ utter rarely (if ever) delivers the kind of complete, truth-evaluable content for sentences envisaged by semantic minimalists. Instead, they hold that hearers need to look to the context of utterance to retrieve a more relevant interpretation of the speaker’s utterance (where ‘relevant’ is understood in a technical sense, as the interpretation which yields the most cognitive effects for the least processing effort). So, returning to our example above, the content recoverable just on the basis of word meaning and structure alone (*Bishma has eaten*) is held to be either incomplete (it can’t be assessed as true or false as it stands) or irrelevant. Thus, hearers are required to recover a more relevant, enriched content, such as *Bishma has eaten recently*.

The contextualist idea that we don’t need to be concerned with the claims made by sentences, independently of what they convey in particular communicative contexts, is reinforced by recognising that advocates of semantic minimalism face two significant problems:

- i. If minimal content is what we get by paying attention to the literal meaning of words (and the way they are put together), then, in order to grasp the minimal content of a sentence we first need to know the literal content of the words it contains. Yet determining the literal, standing meaning of a word (and indeed whether such a thing exists at all) turns out to be a vexed issue.
- ii. As noted above, advocates of minimal semantics accept that minimal content is not generally the content which hearers are concerned with recovering in normal conversational exchanges. Faced with Bishma’s utterance, all parties agree that it is far more likely that hearers will focus on the explicature or implicature content, rather than the minimal claim that *Bishma has eaten something at some point in the past*. Yet why, the objection then is, should we bother positing minimal content at all, if it doesn’t play any important role in communicative exchanges?

I explore objection (i) – that minimal sentence level content is impossible since there is no such thing as minimal word meaning – in the next two sections but conclude that there is a workable notion of minimal word meaning which advocates of semantic minimalism can appeal to. Finally, (in §5) I turn to objection (ii), arguing that an answer to it can be found by reflecting on our practices around assigning and assuming linguistic liability.

3. The Problems of Word Meaning

Intuitively, the question of how we know the ordinary meaning of a word seems easy to answer: if someone is a competent user of a language, we might expect that they can just reflect on what they know about their language (provided we are asking about a word they know) in order to deliver an account of the literal meaning of the word. After all, I relied on this kind of intuitive knowledge of word meaning when I suggested above that ‘I’ refers to the speaker and when I simply assumed that readers would be able to supply the meaning of ‘eaten’ or ‘lunch’. However, just as with our grasp of what is said by a whole utterance, on reflection answering questions about word meaning seems tricky. For instance, take ‘lunch’: roughly speaking, ‘lunch’ refers to a meal taken between breakfast and dinner. But can a meal eaten at 15:00 (or even later) still count as ‘lunch’? And what if someone just eats a packet of crisps or has a drink (is a ‘liquid lunch’ still a lunch)? And even if we all agree that a sandwich eaten at 13:00 does count as ‘lunch’, we then need to know what counts as a ‘sandwich’? If Bishma told Abbie he wanted a sandwich and was then presented with a single piece of bread with cheese on top, would his request have been satisfied? What about if he is given an egg partially between two crackers? (See Searle, 1980.) Prima facie we might expect that a grasp of word meaning should allow us to settle questions about when to apply or to withhold a given expression, but it turns out that settling these kinds of questions is far from straightforward.

One attractive response here would be to object that the kinds of questions posed above about ‘lunch’ or ‘sandwich’ are unnecessarily picky and pedantic (the kind of thing only a philosopher with too much time on their hands might worry about), with little or no practical relevance. In fact, however, we don’t have to look very far to see that these kinds of disputes about word meaning, and about what speakers can or cannot be taken to have asserted, can have serious practical repercussions. To see this, I want to briefly survey two legal disputes from the UK.

3.1 *The meaning of ‘widow’*

The first example revolves around a piece of tax legislation: Section 262 of the Income and Corporation Taxes Act 1988 stated that, in certain circumstances, ‘[a] *widow* shall be entitled [...] to an income tax reduction [...]’. The circumstances in which the allowance came

into effect, and the nature of the allowance itself, are not important here. Instead what matters for the current discussion is to whom the legislation refers. The problems about interpreting Section 262 first came to light when a widower, one Mr Crossland, brought a complaint under the European Commission of Human Rights, alleging breach of Article 14 (discrimination on the basis of sex). The UK tax authority (HMRC) reached an out-of-court settlement with Mr Crossland (paying him £572, the amount of the requested tax relief in his case). Subsequently, however, another widower, one Mr Wilkinson, brought a similar complaint but on this occasion HMRC rejected the claim. They agreed with the complainant that Section 262 was in breach of European law outlawing discrimination on the basis of sex, but they argued that HMRC was nevertheless warranted in enacting it as a matter of statute (arguing that if the UK government wished its tax policy to be compliant with European Human Rights regulation the government itself would have to alter the offending legislation, since it was not within HMRC's powers to contravene extant UK legislation). Mr Wilkinson appealed against this decision and his case was ultimately considered at the High Court. Part of Mr Wilkinson's case at the High Court trial was that Section 262 should in fact be read as applying to him directly, since the term 'widow' in this context should be read as meaning *widows and widowers* due to Section 6 of the Interpretation Act 1978, which states that 'unless the contrary intention appears [...] words importing the feminine gender include the masculine'.

The High Court, however, rejected all aspects of Mr Wilkinson's appeal, including this interpretative claim. The Court pointed out that draftsmen had no difficulty in finding gender neutral terms elsewhere, where there was a clear intention that passages should have the more inclusive reading. Furthermore, other sections of the Income and Corporation Taxes Act 1988 which were intended to apply to both widows and widowers explicitly stated this. Given these facts, Lord Hoffman (for the Court) stated that '[T]here is no way in which any reasonable reader could understand the word "widow" to refer to the more general concept of a surviving spouse'.² Section 262, then, the Court decided, had to be read in a way which placed it in contravention of European law (a fact which led to it subsequently being revoked).

² Wilkinson, R (on the application of) v. Inland Revenue [2005] UKHL 30 (5 May 2005), <http://www.bailii.org/uk/cases/UKHL/2005/30.html>.

What Is It to Be Responsible for What You Say?

3.2 The meaning of 'wife or husband'

The background to our second case, from 2002 (so pre-dating the legalisation of gay marriage in the UK), is that a landlord, Ghaidan, was granted repossession of a flat on the death of the tenancy holder (Wallwyn-James). This repossession order was granted even though the flat was occupied by Wallwyn-James' long-term homosexual partner, Godin-Mendoza. Had Wallwyn-James been survived by a partner of the opposite sex repossession would not have been ordered, as the right to assume a tenancy of those living with a deceased tenancy holder 'as wife or husband' was protected in law. Godin-Mendoza appealed his eviction, arguing that a ruling that homosexual relationships were not covered by UK tenancy laws was in breach of European Conventions. The High Court supported Godin-Mendoza's appeal, ruling that the key phrase in tenancy law – 'as his or her wife or husband' – should be taken to mean 'a relationship of social and sexual intimacy exemplified by, but not limited to, the heterosexual relationship of husband and wife'. In this case, then, the court ruled that the phrase 'wife or husband' was to be understood in the wide sense of *anyone performing the role typical of a wife or husband*.

At least *prima facie* this ruling seems in tension with our first case, where the judgement was that 'widow' could not be understood in the wider sense of 'surviving spouse', but, regardless of this apparent tension, what both cases make clear is that questions concerning the literal meaning of our words can have significant practical repercussions (indeed, as we will see below in the case of Derek Bentley, sometimes they can have the most dire of real-world repercussions).³ So, regardless of the prospects for the philosophical approach of semantic minimalism, it seems that we need a notion of ordinary, literal word meaning (to help settle disputes like those surveyed in this section) and thus it would be good to have a firmer grip on how decisions about the ordinary meaning of an expression are, in fact, made. I explore this issue in the next section.

4. Methods for Determining the Literal Meaning of a Word

The idea that words carry ordinary, literal meanings seems appealing (as it is needed to settle disputes like those canvassed in the previous section, as well as for other reasons, such as explaining how children

³ For an argument that the two verdicts are not in fact in conflict, see Hoffman's judgement on the case discussed in section 3.1.

learn a language, where the knowledge they acquire allows them to express and comprehend an indefinite number of novel word combinations on the basis of limited cognitive resources, see Borg, 2004). However, even if we assume that words do carry this kind of content, there remains an epistemic question about how language users come to know or can decide what the literal meaning of a word actually is. The most common methods for answering questions of word meaning are (as alluded to above) *subjective*: we simply ask what, as a speaker, we intended to mean by uttering some word or, on the other hand, we consider what we, as a hearer, take to be conveyed by the word the speaker produced (as Lord Hoffman apparently did when he pronounced on the meaning of ‘widow’). Subjective approaches are by far the most common route for settling questions of word meaning but they can be problematic. For whenever the meaning of a word is contested and intuitions about word meaning diverge (as when we ask, ‘What constitutes “lunch”?’ or ‘Can “widow” mean *surviving spouse*?’) purely subjective routes leave us with no way to resolve disputes. A further worry is that subjective approaches face the challenge of ‘humpty dumpticism’ – so-called in recognition of Lewis Carroll’s character of Humpty Dumpty in *Through the Looking Glass*, who stated that ‘When I use a word it means just what I choose it to mean – neither more nor less’. Yet, clearly, we don’t want speakers to be the final arbiters of what their words mean in this way (no matter how much I want to, I can’t mean *dog* simply by saying ‘cat’).

In difficult or contested cases, then, it is common to appeal to more objective arbiters of word meaning. For instance, in both ordinary and legal cases, it is common for parties to consult dictionaries, where the definitions of meaning given in these works are taken to be authoritative. Once again, however, it seems that dictionary definitions can’t be the whole story. For one thing, definitions may differ across different dictionaries (meaning that, in legal disputes, defence and prosecution teams often each have their own preferred dictionary). Secondly, dictionaries are unlikely to settle the kind of fine-grained interpretative issues that we saw arise in §3 (a dictionary is unlikely to tell us, for instance, whether a packet of crisps eaten at 15:00 falls within the extension of ‘lunch’).

Recognition of these limitations has led researchers to explore further kinds of objective approaches. For instance, so-called ‘experimental philosophy’ has investigated the use of experimental techniques to help identify the ordinary meaning of words, such as giving speakers questionnaires probing their views on the meaning of a target word or showing them short vignettes and asking whether a

What Is It to Be Responsible for What You Say?

target word would apply in the described situation or not (and then providing statistical analyses of the answers provided); see, e.g., Tobia (2020).⁴ However, experimental techniques are often not practicable since they are time-consuming and expensive. Furthermore, a significant degree of expertise is required to construct experimental materials that properly probe word meaning without introducing confounds. An alternative objective method which has been growing in popularity in recent years (particularly as a tool for assessing the ordinary meaning of legal terms, see Mouritsen 2012; Lee & Mouritsen 2017; 2021; Gries, 2020) is Corpus Linguistics. Corpus Linguistics studies patterns of usage in large datasets of digitised texts, using specialised automatic processing software (e.g., Sketch Engine, WordSmith Tools) to reveal these patterns. Typical analyses include:

- i. Frequency and keyword analysis.
- ii. Collocations (words that appear together with the target word at above chance levels).
- iii. Concordances (sample lines of text containing the target word).

To close our exploration of how we might uncover minimal word meaning, then, I want to look in a little more detail at how this kind of approach might help us get a grip on word meaning in a difficult case.

4.1 *A corpus linguistics case study: 'hate speech'*

The term 'hate speech' originated with legal theorists in the 1980s (primarily in the context of discussions about South Africa) to categorise harmful racist utterances. Many jurisdictions now have hate speech laws (though not all explicitly use the term); e.g. the UK Public Order Act 1986 – commonly considered a piece of hate speech legislation – states that 'A person who uses threatening, abusive or insulting words or behaviour, or displays any written material which is threatening, abusive or insulting, is guilty of an offence if—

- (a) he intends thereby to stir up racial hatred, or
- (b) having regard to all the circumstances racial hatred is likely to be stirred up thereby'.

⁴ Although experimental philosophy as a school of thought is a relatively new approach, this kind of approach to word meaning has its roots in earlier work, in particular Austin's 'Ordinary Language' philosophy can be thought of as an early precursor (see Hansen, 2018).

The expression ‘hate speech’, however, is now also widely used in public discourse (Brown, 2017, p. 424). Yet, this raises the question of whether the meaning of the term is the same in both legal and non-legal contexts. This question is important for a number of reasons: for instance, according to many theories of legal interpretation, the meaning of a legal term is a function of its ordinary meaning, thus grasping the ordinary meaning of ‘hate speech’ matters for questions of legal interpretation. Secondly, a key principle of a democratic system is that of ‘fair notice’, whereby citizens must know, or be able to easily find out, what the law demands of them. This requires that, as far as possible, words are used in legal contexts with their ordinary meanings. Finally, hate speech legislation should not lead to self-censorship of legitimate speech (Howard, 2019; 2021), i.e., it should avoid having a ‘chilling effect’ on legitimate speech. Again, this requires that the legal meaning of ‘hate speech’ coincides with its ordinary meaning. So, it is important that we are clear on the ordinary meaning of ‘hate speech’ and one recent suggestion is that Corpus Linguistics might help advance understanding on this matter.

To assess this claim, Lepoutre *et al.* (2023) ran a pilot corpus study to look at how ‘hate speech’ is used outside the legal realm. The study relied on two corpora:

1. A general corpus (English Web 2020: enTenTen20): containing c. 38 billion words, drawn from English language internet domains 2019–2021. Main genres contained are blogs, discussions, and news articles, with some limited representation of legal sources.
2. A specialised ‘hate speech’ corpus: 255 news reports about hate speech related events, consisting of 164,183 words (retrieved from the Nexis database Lexis Library News, dating from 1990 to 2021). A range of UK media outlets, including ‘serious press’ and tabloids, national newspapers (e.g., *The Daily Mail*) and smaller regional ones (e.g., *The Belfast Telegraph*) were included.

Within (1) and (2), Lepoutre *et al.* assessed:

- frequency (how often ‘hate speech’ and related terms appeared)
- collocations (what expressions ‘hate speech’ appeared alongside)
- concordances (sample lines of text containing the term ‘hate speech’)

The analyses revealed that there has been a *huge* growth in the use of ‘hate speech’ terminology; e.g., in the Lexis Library News database

What Is It to Be Responsible for What You Say?

for British newspapers, in 1990–1995 there were 46 articles mentioning ‘hate speech’, in 2011–2015 there were 2753 articles, while in 2016–2020 this had shot up to 8024 articles. However, the phenomenal spread of the term was not the only finding of note. For instance, results showed that the public understanding of ‘hate speech’ is *wider* than the legal definition in the UK (e.g., it includes hateful speech based on gender, which UK law does not classify as ‘hate speech’). On the other hand, however, the public use of ‘hate speech’ also seemed *narrower* than the legal definition in other respects (e.g. there was no evidence of hateful speech by a powerless minority against a powerful majority being classified as ‘hate speech’, although such speech could fall under the legal definition). Lepoutre *et al.* also found that the ordinary meaning of ‘hate speech’ involves more than ‘offensive speech’, with ordinary speakers often linking it to *incitement*. It seems that these empirical findings could play an ameliorative role. For instance, recognising that the public already classify problematic gender-based speech as ‘hate speech’ supports recent moves by the UK government to extend the legal definition in this direction, while the realisation that ordinary speakers do not simply classify all offensive speech as ‘hate speech’ helps to mitigate concerns about the chilling effect of hate speech legislation.

To conclude, then, whilst it seems that no single method for investigating ordinary word meaning is likely to provide a complete answer to questions about the ordinary meaning of our words, in contested cases it does seem that objective methods, like corpus analysis, may help to improve our understanding (Lee & Mouritsen, 2021, p. 358). When questions about the standing meaning of a word are raised, a mixture of subjective and objective methods should be used in order to triangulate on the most plausible analysis of the expression’s ordinary meaning. Given this range of methods, though, I want to suggest that the first worry raised for minimal semantics (in §2) – that we don’t have a way to determine literal, or minimal, word meanings – can be rejected.

However, this still leaves the second worry pending. Recall that the second objection to minimal content was that the literal, context-independent meaning of sentences was irrelevant as far as questions of communication were concerned. Even if there is a case to be made for needing to know about the ordinary, literal meaning of *words* (in order to resolve problem cases like the ones considered in §3), it is unclear that this will provide any support for the semantic minimalist’s claim that we need to be concerned about minimal *sentence* contents (the contents that emerge just from putting words with those meanings together in that particular grammatical structure). If the

contents which really matter for communicative exchanges are the contextually enriched contents of utterances (e.g., in our example of ‘I’ve eaten’ from §2, contents such as *I’ve eaten recently* or *I don’t want to have lunch*), why, the objection remains, should we bother trying to discern some more context-free content? This brings us back to the question of the explanatory worth of minimal content and the answer I want to give is that the minimal contents generated by the literal meaning of words (plus sentential structure or grammar) matter because they are needed to answer certain questions that arise about linguistic liability.

5. Linguistic Liability

As noted at the outset, it seems that we have a practice of taking speakers to be responsible for the truth of what they assert. That is to say, speakers are generally held to have ‘linguistic liability’ for the contents they express – they can be held liable for what they say. However, as argued in Borg (2019) and Borg & Connolly (2022), it seems that our practices of acquiring and assigning linguistic liability are multifaceted. Specifically, it seems that there are two distinct varieties of linguistic liability:

- i. Strict linguistic liability: a binary notion whereby a speaker either is or is not held liable for the *literal content* of the sentence, s, she uttered.
- ii. Conversational linguistic liability: a matter of degree whereby a speaker is held more or less liable for some *non-literal content* via her utterance of s.

Compare two utterances of the sentence ‘The train leaves at 11am’: in the first scenario, imagine that a friend in Oxford has just asked me if I’ll join him for lunch and I reply by saying ‘I’m going to London today. The train leaves at 11am’. In the second scenario, imagine that I’ve phoned National Rail inquiries to ask about the next train to London today and an employee tells me ‘The train leaves at 11am’. It seems that the kinds of liability operative in these two situations differs. In the first scenario, it seems that only conversational liability is in play – I’m trying to say that my train leaves at a time which makes having lunch together impossible, thus my hearer is unlikely to feel aggrieved if he finds out that the train I’m catching actually leaves at 10.56am or 11.03, or any other time roughly around 11am. On the other hand, however, when I speak to the rail employee, the standards of liability are much stricter. This time I’m entitled to hold the

What Is It to Be Responsible for What You Say?

speaker to the truth of what they literally say – if I arrive at the platform at 10.57 only to see the train departing, as it was actually scheduled for departure at 10.56, it seems I have every reason to feel aggrieved. In the above chat amongst friends about their lunch plans, liability goes along with the contextually enriched content, but in the exchange with the rail worker liability rests with literal content.

As with our earlier discussion of literal word meanings, this philosophical distinction is of more than merely academic interest. It is relevant both in ordinary exchanges like the ones above and in higher stakes disputes. For instance, consider the infamous 2012 ‘Twitter Joke Trial’ (Chambers v Director of Public Prosecutions 2012). The background facts of this case were that a 26-year-old UK citizen, Paul Chambers, on learning that an airport from which he was due to travel was closed due to heavy snowfall, responded on Twitter with: *‘Crap! Robin Hood airport is closed. You’ve got a week and a bit to get your s**t together otherwise I’m blowing the airport sky high!’*. Five days later, the duty manager of airport security read the tweet and communicated it to the police. Chambers was duly charged with, and subsequently convicted in a Magistrates Court of, ‘sending a message of a menacing character’ contrary to the Communications Act 2003. Chambers appealed this conviction, which was at first upheld in the Crown Court but eventually quashed at the High Court. At each stage of this case, we can see hearers’ sensitivity to strict versus conversational liability. For instance, the airport security had a clear protocol for terrorist threats which required them to report any such material to the Ministry of Defence. The fact that they did not report Chambers’ tweet in this way, when the literal content of ‘I’m going to blow the airport sky high’ would undoubtedly qualify as such a threat, shows that they did not apply standards of strict liability: they did not take Chambers to be liable for the literal content of his utterance. On the other hand, the fact that they still reported the tweet to the police shows that they thought Chambers could be held strongly conversationally liable for some non-literal threatening content, a judgement that was upheld by both the Magistrate and the Crown courts (where judges asserted that the tweet was ‘clearly menacing’). The High Court, on the other hand, whilst agreeing that Chambers was not liable for the strict literal content of his utterance, also judged him not to be conversationally liable for any threatening content, stating that ‘the more one reflects on it, the clearer it becomes that this message did not represent a terrorist threat, or indeed any other form of threat [...]. The language and punctuation are

inconsistent with the writer intending it to be or to be taken as a serious warning'. If we want to explain the basis of these various different judgements in the Twitter Trial we need both to recognise that different kinds of linguistic liability can be assigned and that one kind of liability judgement tracks the sort of strict, literal content proposed by semantic minimalists.

Finally, it seems that sometimes the question of whether to attribute strict linguistic liability to a speaker or instead to hold them only to a more pragmatic (including, in the case to be discussed, idiomatic) content can have the most serious of consequences. One such case concerns the prosecution (in the UK in 1953) of Derek Bentley for the murder of a policeman. Bentley, who was aged just 18 at the time of the offence, was involved in an attempt to rob a warehouse with an accomplice, Christopher Craig (aged 16). However, police were alerted during the attempted break in and the first policeman who arrived on the scene (Frederick Fairfax) was able to grab hold of Bentley, at which point Craig drew the gun he was carrying. All parties agreed that Fairfax shouted 'Give me the gun lad', at which point Bentley allegedly replied with a key phrase: 'Let him have it Chris'. Soon after Fairfax's utterance, Craig fired the gun, hitting and injuring Fairfax. A short time later he shot and killed another policeman, Sidney Miles.

At the criminal case, Bentley's legal team denied that Bentley had ever uttered the words 'Let him have it'. However, they also argued that, even if Bentley were believed to have uttered the phrase in question, the meaning of the utterance in this context would have been *Let the policeman have the firearm Chris*. On the contrary, the prosecution team argued that the correct interpretation was an incitement to violence: *Shoot him Chris*. In the end, both Craig and Bentley were found guilty of the murder of Sidney Miles under the legal rule of Joint Enterprise. Craig, as he was under 18, was sentenced to be detained at Her Majesty's Pleasure (eventually being released ten years later). Bentley, on the other hand, was sentenced to death. Bentley made an unsuccessful appeal and, despite a public outcry, he was executed by hanging in January 1953. Following a 40-year campaign by his family, however, Bentley received a posthumous pardon in 1993 and his conviction was finally overturned in 1998.

Bentley's case is clearly extremely disturbing, but it does throw into sharp relief the importance that settling questions of linguistic liability can have. Should Bentley have been taken to be responsible for the conversational, idiomatic content (*Shoot him*) or instead for a literal interpretation (along the lines of *give it – the salient object – to the policeman*)? And what factors could, or should, jurors have

What Is It to Be Responsible for What You Say?

looked to when considering this question? While the conditions which dictate the appropriateness of either strict or conversational liability attributions are no doubt highly complex and hard to state (see Borg & Connolly, 2023), and we might well query whether the jury in this case got things right, what cases like this one demonstrate is the need for ordinary interlocutors to clearly distinguish different kinds of linguistic liability and for researchers to work to clarify how the standards of linguistic liability get decided in different contexts.

Along with more everyday practices (such as understanding the differences between lying, misleading, and merely loose talk, see Borg, 2019) these legal cases demonstrate that:

- Ordinary speakers are adept at making different kinds of liability judgements, where these judgements are sensitive to different kinds of content.
- Some judgements of linguistic liability are strict and strict judgements require a grasp of minimal content.

Contra the irrelevancy objection to minimal semantics (raised as objection (2) in §2), then, it seems clear that, from both a practical and a theoretical perspective, we do need to recognise minimal content. While it may not capture the content that is most commonly at play in communicative exchanges, we need to keep track of the content a speaker expresses simply on the basis of the literal meaning of their words and the way those words are put together because, at least sometimes, this is the content which speakers acquire liability for. Our practices around assigning and accepting linguistic liability, then, provide a clear explanatory role for minimal content. When Donald Trump said that the crowd for his inauguration stretched ‘all the way back to the Washington Monument’ the journalists reporting him held him to a standard of strict linguistic liability, holding him responsible for the truth of what he literally said, a content shown to be false by photographic evidence that the crowd stopped far short of the Monument.

6. Conclusion

Generally, we hold speakers to be responsible for the truth of what they say. This practice is crucial to the role that language plays in our lives, for without it we would be unable to use the testimony of others to direct our interactions with the world and to further our understanding. However, although from a practical perspective we usually have little problem fixing on the contents to which we

should hold speakers committed, from a theoretical perspective things are less clear. For to hold someone responsible for what they say, we first need to know:

1. The content of what they say.
2. The sense of responsibility in play.

The most common methods for assessing (1) are subjective: generally, it is intuitively clear what a given speaker should be held to have said in a given context. As we have seen, however, a simple appeal to subjective intuitions may be inadequate in difficult or contested cases. Once we start wondering about what exactly a word means, and what worldly states of affairs it does or doesn't apply to, it can begin to seem that we lack appropriate resources for settling these questions. In response, I suggested (§4) that we should investigate the options for more objective routes to word meaning, looking not just at what one or two interlocutors might suggest about word meaning but surveying assumptions about meaning across a population (e.g., via running experiments about ordinary meaning or conducting corpus analyses). Although no one method is likely to prove decisive, by gathering evidence about what a wide sweep of ordinary language users think about the meaning of some contested term, we can help to balance out the idiosyncrasies and peculiarities of individual language users, arriving at robust analyses of the ordinary, standing meaning of our words.

Furthermore, getting clear on ordinary, literal meaning matters because some perfectly standard practices around assigning liability *depend* on this kind of content. We sometimes take speakers at their word: we apply standards of *strict linguistic liability* and (contra the contextualist objection that minimal semantic contents have no explanatory role to play) this requires hearers to be sensitive to the content that emerges from the literal meaning of words plus the grammatical structure of sentences. I suggest, then, that understanding the different ways in which someone can be held responsible for what they say reveals the explanatory need for minimal content (even whilst we recognise that in the cut and thrust of communication it is often non-minimal content which is to the fore).

Acknowledgement

Research for this work was supported by a Leverhulme Trust Major Research Fellowship MRF-2019-031.

References

- Kent Bach, 'The Top Ten Misconceptions about Implicature', in Betty Birner and Gregory Ward (eds), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn* (Amsterdam: John Benjamins, 2006), 21–30.
- Emma Borg and Patrick Connolly, 'Strict and Conversational Linguistic Liability: Some Empirical Observations', in Ernest Lepore and David Sosa (eds), *Oxford Studies in Philosophy of Language* (Oxford: Oxford University Press, 2022), 1–22.
- Emma Borg, *Minimal Semantics* (Oxford: Oxford University Press, 2004).
- Emma Borg, *Pursuing Meaning* (Oxford: Oxford University Press, 2012).
- Emma Borg, 'Explanatory Roles for Minimal Content', *Noûs*, 53:3 (2019), 513–39.
- Alexander Brown, 'What is Hate Speech? Part 1: The Myth of Hate', *Law and Philosophy*, 36:4 (2017), 419–68.
- Robyn Carston, *Thoughts and Utterances* (Oxford: Blackwell, 2002).
- Robyn Carston, 'Review of *Minimal Semantics*', *Mind & Language*, 23:3 (2008), 359–67.
- Paul Grice, *Studies in the Way of Words* (Cambridge, MA: Harvard University Press, 1989).
- Stefan Gries, 'Corpora and Legal Interpretation: Corpus Approaches to Ordinary Meaning in Legal Interpretation', in Malcolm Coulthard and Alison Johnson (eds), *Routledge Handbook of Forensic Linguistics*, 2nd edition (London: Routledge, 2020), 628–43.
- Nathaniel Hansen, '"Nobody Would Really Talk That Way!" The Critical Project in Contemporary Ordinary Language Philosophy', *Synthese*, 197:6 (2018), 2433–64.
- Jeffrey Howard, 'Dangerous Speech', *Philosophy & Public Affairs*, 47:2 (2019), 208–54.
- Jeffrey Howard, 'Terror, Hate, and the Demands of Counterspeech', *British Journal of Political Science*, 51:3 (2021), 924–39.
- Thomas Lee and Stephen Mouritsen, 'Judging Ordinary Meaning', *Yale Law Journal*, 127:4 (2017), 788–879.
- Thomas Lee and Stephen Mouritsen, 'The Corpus and the Critics', *University Chicago Law Review*, 88:2 (2021), 275–366.

- Maxime Lepoutre, Sara Vilar-Lluch, Emma Borg, and Nat Hansen, 'What is Hate Speech? The Case for a Corpus Approach', *Criminal Law and Philosophy*, (2023).
- Stephen Levinson, *Presumptive Meanings: The Theory of Generalized Conversational Implicature* (Cambridge, MA: MIT Press, 2000).
- Stephen Mouritsen, 'Hard Cases and Hard Data', *Columbia Science & Technology Law Review*, 13:1 (2012), 156–205.
- John Searle, 'The Background of Meaning', in John Searle, Ferenc Keifer, and Manfred Bierwisch (eds), *Speech Act Theory and Pragmatics* (Dordrecht: Reidel, 1980), 221–32.
- Dan Sperber and Deirdre Wilson, *Relevance: Communication and Cognition* (Oxford: Blackwell, 1986).
- Kevin Tobia, 'Testing Ordinary Meaning', *Harvard Law Review*, 134:2 (2020), 726–806.
- Timothy Williamson, 'Knowing and Asserting', *The Philosophical Review*, 105:4 (1996), 489–523.