Language and Law / Linguagem e Direito, *Vol. 11*(2), 2024, pp. 24-54
DOI: doi.org/10.21747/21833745/lanlaw/11_2a8

Offence-Presumptive Terms: A Troubling Category in Linguistic Offensiveness†

Chris Heffer [1]

[1] *Cardiff University*, HefferC2@cardiff.ac.uk

Abstract
This paper argues that there is a set of words, word senses and phrases that are not inherently offensive but that can lead the hearer to presume that they are offensive. The archetypal, though now somewhat problematic, case is the adjective 'niggardly' ('stingy', 'parsimonious'), but I also discuss two cases based on forensic casework: the metaphorical sense of the verb 'kneecap' (bring to its knees); and the now aging catchphrase 'never mind the quality, feel the width' (go for quantity over quality, used ironically). These offence-presumptive terms are a subcategory of Low-Occurrence Ordinary Terms (LOOTs). LOOTs are ordinary enough for the speaker to presume they will be understood and for the hearer to presume that they know or can infer the meaning, while at the same time infrequent enough not to be familiar to many hearers. Crucially, offense-presumptive LOOTs have lexical 'neighbours' in our mental lexicons that are either attested as, or are perceived to be, offensive. I argue that the presumption of offensiveness comes about because hearers apply autonomous intuitive reasoning to the unknown word, word sense or phrase. While in many cases of lexical inferencing, such intuitive reasoning can lead us to make a successful guess about the meaning of a word, in the case of offence-presumptive terms, the hearer is led astray. Moreover, in raising a claim of linguistic offence, claimants prime HR personnel to activate in their own minds an offensive frame for the term. I consider the implications of this novel category for forensic lexical analysis, the judgment of linguistic offence and the ethics of interpersonal communication.
Keywords: *offensive language, lexical neighbours, forensic lexis, dual-process reasoning, lexical inferencing, phonological similarity,* niggardly*, N-word.*

Resumo
Este artigo defende que existe um conjunto de palavras, significados e orações que não são inerentemente ofensivos, mas que podem levar o ouvinte a presumir que o são.

† **Content Note** Following standard linguistic practice in the area of offensive language, this article mentions and quotes taboo words including slurs.

O caso prototípico, embora agora um pouco problemático, é o adjetivo "niggardly" ("mesquinho", "avarento"). Além deste, discuto também outros dois casos baseados em trabalhos forenses: o sentido metafórico do verbo "kneecap" ("ajoelhar-se"); e a expressão agora antiquada "never mind the quality, feel the width" ("prefira a quantidade à qualidade", usada ironicamente). Esses termos presumivelmente ofensivos são uma subcategoria dos Termos Comuns de Baixa Ocorrência (LOOTs). Os LOOTs são comuns o suficiente para que o falante presuma que serão compreendidos e para que o ouvinte presuma que conhece ou pode inferir o seu significado; simultaneamente, também são pouco frequentes para não serem familiares a muitos ouvintes. Na sua essência, os LOOTs, que podem ser considerados ofensivos, têm "vizinhos" lexicais nos nossos léxicos mentais que são comprovadamente ofensivos ou são percebidos como tal. Discuto aqui que a presunção de ofensividade surge porque os ouvintes aplicam um raciocínio intuitivo autónomo à palavra, sentido da palavra ou oração desconhecida. Embora em muitos casos de inferência lexical, esse raciocínio intuitivo possa levar-nos a adivinhar com sucesso o significado de uma palavra, no caso de termos que podem ser considerados ofensivos, o ouvinte é induzido em erro. Além disso, ao apresentar uma alegação de ofensa linguística, os queixosos levam o pessoal de RH a ativar nas suas próprias mentes uma interpretação ofensiva para o termo. Considero ainda as implicações dessa nova categoria para a análise lexical forense, o julgamento da ofensa linguística e a ética da comunicação interpessoal.
Palavras-chave: *Linguagem ofensiva, vizinhos lexicais, léxico forense, tarefa dupla de raciocínio, inferência lexical, semelhança fonológica, "niggardly", N-word.*

1. Introduction

This article develops a novel argument relating to a particularly troublesome category of claimed linguistic offensiveness: the case where a claimant is deeply offended on hearing a word or phrase that is neither objectively offensive from a linguistic standpoint nor intended by the speaker to be offensive. The archetypal (though now highly controversial) case is that of the adjective *niggardly* ('stingy'), which is phonologically similar to the N-word but etymologically and semantically completely unrelated to it. In a famous case in 1999, the director of a municipal agency in Washington, David Howard, told his staff that due to severe financial constraints, he would need to be 'niggardly' with the budget. A couple of Howard's Black staff 'began a whispering campaign that blossomed into a public outcry' (Kennedy, 2002, p. 94) which led to the mayor accepting Howard's resignation on the grounds that he had shown 'poor judgement.' Yet Howard had used the word in its most common context of use (indicating financial constraints) and there was no suggestion that he intended to be offensive. Nor, at the time, was there extensive media discussion of the word of which he could have been aware. While the first

public complaint about the use of *niggardly* appears to have been by a Boston reader in 1995 complaining about The *Economist*'s use of the word in 'productivity growth averaged a niggardly 0.8% a year', the UK-based magazine simply treated the complaint with condescending amusement (Derbyshire, 2002). Reactions to the Howard scandal were as deeply polarised then as controversies about offensive language are today and showed very little understanding about the nature of language and communication. On the 'offence' side, one columnist asked whether Howard could possibly not have noticed that 'he had to pass "nigger" before he could get to the "dly"' (Dickerson, 1999). But this is simply not how we process language: we do not have to 'pass' *cunt* when uttering *country* and a large proportion of informants who are asked what Rice Krispies are made of are unable to respond 'rice' (Wray & Staczek, 2025) despite having to 'pass' the word when naming the brand. Instead, we store words and phrases as single meaning units ('stingy' or 'brand of breakfast cereal') and, even in the case of homonyms, we automatically select the word or sense that is most appropriate in the context. This is why it is not unusual in Wales to see adverts for faggots (Welsh meatballs) such as the following:

> **Faggots** are an everyday favourite here in Pembrokeshire, so we've created what we see as the perfect Pembrokeshire **faggot** – see what you think! (Gwaum Valley Meats at https://gwaunvalleymeats.co.uk/product/gwaun-valley-faggots. Accessed 28/6/25)

Just as the writer of this advert almost certainly was not thinking of the gay slur when writing *faggot*, so Howard was most likely unaware of the phonological proximity to the N-word when uttering the contextually appropriate word *niggardly*, and one wonders how many teachers notice the proximity of *snigger* to *nigger* when telling pupils to 'stop sniggering at the back of the class'. At the same time, the 'defence' reactions also showed ignorance of language. One columnist described Howard as the victim of 'linguistic lynching' who got fired because 'some people in public employ were morons who … didn't know how to use a dictionary' (Snow, 1999). Yet offensiveness does not lie in the dictionary meaning of a word but in how it is used in a specific communicational context (Heffer, in prep). Chris Rock's frequent use of the N-word in his comedy routines presumably does not offend his usual audiences while a White worker might pointedly use *niggardly* to rile a Black co-worker who she knows is offended by the word. Claimants in such cases are often dismissed as hypersensitive or vexatious. Yet, the hurt can be real. Hearing perceived taboo words can give an emotional jolt that has been measured through skin conductance (Dinn & Harris, 2000) and effect on the amygdala (Hansen, McMahon, & De Zubicaray, 2019). So it is perfectly possible for a claimant to feel genuinely harmed by a perceived linguistic offence even if the offence was entirely unintentional and not objectively warranted by the evidence.

Within linguistics, the issue of offence-adjacent 'innocent' words and senses being perceived as offensive tends to be viewed with respect to processes of tabooification in the development of the language as a whole (Allan & Burridge, 2006; Pinker, 2007) and linked to 'verbal hygiene' (Cameron, 1995), or our deliberate attempts to 'clean up' the language. Pinker (2007), for example, explains the tabooification of such words in terms of a linguistic version of Gresham's Law ('bad money drives out good'): 'bad words drive good words out of circulation' because people 'often avoid using innocent terms that they fear might be misheard as profanity' (Pinker 2007, p. 333). Allan and Burridge, in

their introduction to taboo language, write that 'Speakers will not risk appearing to use a taboo term when none was intended; therefore, they are quick to drop the homonyms of taboo terms' and that '[e]ven words and phrases that are similar to non-PC terms are avoided' (Allan & Burridge, 2006, p. 102). They conclude that:

> As is so often the case in issues to do with language, it doesn't matter what the linguistic facts suggest. The reality that *niggardly* has absolutely no etymological connections with *nigger* is of no consequence. What really matters is how speakers perceive their language to be, and if people do start connecting words such as *nitty-gritty* and *niggardly* with the N-word, then this will be the kiss of death for these words. (2006, p. 104)

They go on to point out that *fuk* ('sail') and *feck* ('effect') 'had absolutely nothing to do with *fuck*, but that did not save them' (2006, p. 104-105).

The process of tabooification, though, as can now be evidenced through large language corpora, actually tends to be more gradual, piecemeal and reversible than these writers seem to suggest. Even where the process is near complete for one sense or grammatical form, it may not be for another. For example, *cripple* (n) has taken decades to become more or less universally taboo when referred to a disabled person in the US and the UK, but as this use has dramatically declined, the metaphorical use ('Truss *crippled* the economy') has significantly increased (Google NGram Viewer 'cripple') rather than suffering 'the kiss of death' through its association with the noun. *Fuck* has been taboo in published writing for centuries and in public broadcasting for decades (throughout remaining common in private speech) but it is far less taboo today than it was 50 years ago. The process of tabooification can be restricted to region (e.g. *cock* in the US but not the UK), age (e.g. the 'F-word' has different references when uttered by the old (*fuck*) and young (*faggot*)) and many other factors. Twenty years after Allan and Burridge (2006), *nitty-gritty* and *niggardly* have not suffered the kiss of death: they are contested from time to time but ultimately the 'linguistic facts' do seem to matter and each time a complaint is made the public are reminded of those facts. Similarly, as we have seen with *faggot*, speakers are not always so quick to drop the homonyms of taboo terms but, instead, offensive-innocent pairs often co-exist peacefully.

While linguists might appear resigned to an 'inevitable' if (one suspects) 'regrettable' tabooification process of 'innocent' words, a recent strand of thinking in the philosophy of language actively prescribes the tabooification of slurs and any words associated with them. While eradicationists, who argue that all uses of the N-word are hurtful and wrong and so should be prohibited from being used by Blacks any more than by Whites (Kennedy, 2002, p. 126-128), have been around for a long time, Stojnić and Lepore (2025) take this prohibitionism to a new level. In their 'articulation account', a slur has a 'pejorative potential' that 'robustly projects out of environments that normally render meaning inert – including not only quotation and meaning attributions, but even mere displays of slurs' (Stojnić & Lepore, 2025, p. 123). In other words, one cannot articulate a pejorative word even in academic discussion such as this. Worse still, '[p]ejorative potential can easily "infect" expressions whose articulations only incidentally match, or resemble, those of slurs' (Stojnić & Lepore, 2025, p. 123). Accordingly, the authors self-censor their mentioning of these 'infected' expressions and have to rely on awkward circumlocutions instead:

> Recall, once again, the incidents involving the tokening of the English adverb[1] that happens to both orthographically and phonetically resemble the N-word but is otherwise etymologically and semantically unrelated. (Stojnić & Lepore 2025, p. 123)

Such fetishization of language, though, fails to acknowledge that offence lies in a given act of communication rather than in the word itself. In a linguistic study such as this, the taboo word is an object of study, like a sample under a microscope; just as an excised tumour on a slide might look unpleasant but would not generally be considered harmful to the lab technician, a taboo word extracted from its normal contexts of use and presented as a linguistic sample in an academic discussion should not generally be considered harmful to the academic reader interested in taboo language. I follow here, then, the standard linguistic practice of referring to the linguistic objects of study when needed, though I mostly use the 'N-word' euphemism because of the hypersensitivity relating to that term.

Given the heightened emotions excited by the N-word, and anything apparently related to it, I would not tackle this topic were it not for forensic casework I have been engaged in over the past few years which involved disputed terms such as *kneecap* and *never mind the quality, feel the width*, which are unmarked by known historical controversies that might have affected the perception of offence. In such cases, I observed three phenomena in particular. Firstly, the items were both ordinary rather than specialised terms and yet occurred infrequently in the language as a whole. Secondly, despite probably not knowing the intended meaning of the term or sense, the claimant had such a strong presumption that they knew the term was offensive that they were willing to lodge a formal complaint. Thirdly, despite there being no linguistic or circumstantial evidence that the disputed word or phrase was offensive in the context, HR personnel were happy to advance the cases through disciplinary proceedings. Rather than taking sides in the culture war on linguistic offensiveness, though, the argument I develop here presumes a situation where the claimant was genuinely offended but where the speaker had neither intention of causing offence nor could have known that what they said might cause offence.

The argument proceeds in six steps, which will be teased out in the remainder of the article:

1. Since communication is not a conduit, there is always the risk that a hearer will not understand a word or phrase used by the speaker or will interpret it in a way unintended by that speaker;
2. This risk increases markedly in the case of a class of words I propose to call Low Occurrence Ordinary Terms (LOOTs). The ordinariness of a LOOT means both that the speaker will expect the hearer to understand it and that the hearer will expect to be able to infer the meaning if it is not part of their lexicon. But the low occurrence of a LOOT (for an ordinary term) means that many hearers will not be familiar with it and may well misconstrue it;

[1] Discussions of *niggardly* in the media and non-linguistic academic publications tend to erroneously describe it as an adverb, possibly because of the *-ly* derivational suffix. The adverb used to exist, as in 'I hope you did not feel that I had treated you niggardly' (OED 2003: *niggardly* adv), but this 1959 citation is the last attested for the adverb in the OED, which describes the adverb as 'now rare'.

3. Given the ordinariness of a LOOT, when hearers are not familiar with it, they are likely to use autonomous intuitive processes to infer its meaning, such as background contextual inferencing and lexical association, rather than deliberative reflective processes such as morphological or componential analysis or looking up the word in a dictionary. Such 'Type 1' associative reasoning often succeeds in inferring the correct or approximate meaning of the unknown term, while 'Type 2' reflective inferencing is often erroneous;
4. However, where a LOOT is phonologically similar or identical to a recognised offensive term or sense in our mental lexicons, autonomous processing is likely to incorrectly associate the neutral LOOT with the (often more common) offensive term or sense;
5. Furthermore, when Type 1 associative processing mis-leads the hearer towards the offensive lexical neighbour, the lack of conscious reflection involved is likely to mean that the hearer will simply presume that the neutral term is offensive and that the speaker is being offensive and thus not consciously reflect on whether this is actually the case;
6. Finally, by raising a formal claim that the word or phrase is offensive, HR personnel are metalinguistically primed to frame the offence-presumptive term as offensive.

The following five sections tease out the first five steps in the argument outlined above. I then operationalise the notion of 'offence-presumptive term' so that it can be applied systematically to data. Next, I re-analyse two examples from casework that might be considered offence-presumptive terms: the verb *kneecap* and the phrase *never mind the quality, feel the width*. Then I consider the question of metalinguistic priming in the framing of the term as offensive. Finally, I discuss the implications of this novel lexical category in terms of forensic lexical analysis in general, judging linguistic offence in contexts such as workplace harassment and hate speech in particular, and the ethics of interpersonal communication more broadly.

2. Communicational Risk

Were communication a flawless exchange of shared meanings packaged in transparent signs, there would be no problem with miscommunications and misunderstandings. Both the folk linguistic view and early academic approaches to communication shared what might be called a 'Code and Conduit' model of language. The Code consists in lexical signs and grammatical codes that are shared unproblematically among speakers of a language. These shared codes are then meant to be conveyed fluently through a communicational Conduit between interlocutors, as in Saussure's 1983 famous 'talking heads' diagram of the 'speech circuit' or Shannon and Weaver's Bell-(now AT&T)-funded highly influential model of communication (Shannon, 1948) in which communicating was about 'reproducing at one point … a message selected at another point' (Shannon & Weaver, 1964, p. 31).

Over the last fifty years, though, sociolinguistics, linguistic pragmatics and psycholinguistics have troubled this 'conduit metaphor' (Reddy, 1979) of verbal communication. With respect to the Code, while fluent speakers of a language undoubtedly share a common core of linguistic resources consisting in core lexis (e.g. basic verbs, function words) and grammar (e.g. the SVOCA standard word order in English), be-

yond that core there is an enormous amount of individual variation in our knowledge of language. While we talk of 'speech communities' (Gumperz, 1993), 'discourse communities' (Swales, 1990) or 'communities of practice' (Lave & Wenger, 1991), such linguistic communities are sometimes seen by sociolinguists as little more than aggregates of the idiolects of each individual speaker in that community (Hudson, 1996). When we switch focus from the community to the individual, as in forensic authorship attribution, idiolects become paramount (Coulthard, 2004; Grant, 2013). In particular, our individual mental lexicons of words in long-term memory (Jackendoff, 2002) not only differ from others in the words we have stored there, but our individual representations of those words we apparently share with others are unique because of our own particular lifetime experience with those words. Furthermore, the experiences that will have shaped people's understanding of words and grammar in a multicultural and multilingual community are likely to be even more diverse than in more monocultural communities in the past.

Turning to the Conduit aspect, pragmatics has shown that interpretation depends very heavily on context and that speakers can exploit this knowledge by conveying indirect meaning through, for example, implicature and irony (Dynel, 2018; Grice, 1975). But unlike the conduit model of semantic word containers relayed through a communication pipeline, pragmatic meaning is more hazardous. Implicature and irony require 'uptake', or recognition by the hearer of the speaker's communicative intention, and this is by no means guaranteed (L. McDonald, 2020). For example, autistic children and adults are well known to have difficulty distinguishing literal and non-literal language (Happé, 1995). Similarly, psycholinguistics and cognitive neuroscience have come to see language comprehension as context-dependent, predictive and probabilistic (Federmeier, 2007; S. A. McDonald & Shillcock, 2003). The fact that we 'think ahead' when comprehending language has been demonstrated through eye-tracking experiments and through the measurement of Event-Related Potentials (ERP), or brain responses to a specific linguistic stimulus (Federmeier, 2007). But it has also long been known in conversation analysis: micro-analysis of turn-taking (Sacks, Schegloff, & Jefferson, 1974) shows that listeners monitor the trajectory of an ongoing speaker turn for Transition Relevance Places (TRPs) and often try to take the conversational floor once they can predict the end of the speaker's turn (de Ruiter, Mitterer, & Enfield, 2006). However, prediction always comes with risk: the predicted word in lexical comprehension might turn out to be wrong, or listeners might try several times unsuccessfully to take the floor and if they persist without the speaker ceding the floor, it will be interpreted as an unwanted interruption (Kurtić, Brown, & Wells, 2013).

Given the idiolectal nature of language and the predictive nature of language processing, we often find ourselves in a position where we are not familiar with a word or phrase or it seems to be being used in a sense we do not recognise. In other words, the word or word sense is not in our mental lexicons.

3. Low Occurrence Ordinary Terms (LOOTs)

The risk of a hearer misconstruing the meaning of a word or phrase significantly increases in the case of a class of words I call Low Occurrence Ordinary Terms, or LOOTs. The idea of a LOOT has emerged from forensic casework, where I have observed that words and phrases in dispute often belong to this class. There are good reasons for this.

There are two significant aspects to LOOTs that increase the risk of communicational trouble: their ordinariness and their comparatively low occurrence.

'Ordinary Terms' are words, word senses or phrases that would be considered to belong to 'ordinary language' rather than specialist, technical, arcane or archaic language. Ordinary terms can be found in everyday contemporary spoken and written genres such as conversation, public talks, fiction and written and broadcast media. Accordingly, rather than being exclusionary, like professional jargon and circumscribed slang, they are inclusionary in terms of potential audiences. The ordinariness of a LOOT, then, has two consequences:

1. The speaker will expect the hearer to be familiar with the term (Howard would have expected his audience to understand the 'ordinary' adjective *niggardly*); and
2. The hearer will expect to be able to infer its meaning if it is not part of their lexicon (his audience would hear 'be niggardly with the budget' as involving ordinary behaviour rather than some obscure technical process)

These consequences do not follow, on the other hand, if a term is recognised as technical or specialist. The speaker, all things being equal, will not expect the hearer to understand the technical or specialist term unless they are part of the technical or specialist community in which it is used. While professionals sometimes use technical terms to perform expertise and erudition in front of lay audiences (Billig, 2013), they are being exclusionary in so doing because they should know that the audience will not be familiar with such terms. Similarly, if the hearer perceives an unfamiliar term to be technical, they are more likely to accept that they just do not understand it rather than presuming that they understand it or can infer its meaning. Some recognized linguistic cues to technical status include word length (Oppenheimer, 2006) (e.g. *antidisestablishmentarianism*), complex morphology (e.g. *photosynthesis*), certain types of derivational morphemes (e.g. *-ology*, *-itis*), nominalizations (Kies, 1985) (e.g. *facilitation*, *enrichment*) and technical discourse contexts such as scientific or academic publications or technical manuals. One reason why legal language is so problematic is that so many legal terms are disguised as ordinary language terms: *trust*, *person*, *sure* etc.

The other aspect of LOOTs, 'Low Occurrence', needs to be qualified in two respects. Firstly, it is low relative to the ordinary status of the LOOT, not in terms of absolute frequencies in the language. The OED's frequency bands (OED, 2025), based on the Google Books database, are helpful in explaining this. The bands run from 1 (very low frequency) to 8 (very high frequency). Band 1 words (e.g. *abaptiston, zarnich*) are effectively obsolete. Band 2 words (occurring fewer than 0.01 times per million words in contemporary discourse) (e.g. *ennead, abactinal* and *absterge*) are described by the OED as 'almost exclusively terms which are not part of normal discourse and would be unknown to most people' (OED, 2025). The lowest frequency band that might contain ordinary terms, then, is Band 3 (0.01 to 0.1 words per million) (e.g. *amortizable, prelapsarian, agglutinative*), which the OED describes as follows: 'These words are not commonly found in general text types like novels and newspapers, but at the same time they are not overly opaque or obscure.' LOOTs mostly belong to this Band 3 range occurring between 0.01 to 0.1 words per million. However, what characterizes a LOOT in comparison with other words in this band is that, though it does not occur frequently,

it occurs *mostly* in everyday genres such as conversation and the media. The combination of presence in everyday genres but low overall occurrence has the following consequences:

1. The speaker, recognizing its presence in everyday genres, is likely to consider the term 'ordinary' and thus understandable to the hearer;
2. The hearer, though, is likely not to be familiar with the term because of its overall low occurrence.

The second qualification is that while the occurrence is low with respect to the language as a whole, as represented by large reference corpora such as the TenTen family (Jakubíček, Kilgarriff, Kovář, Rychlý, & Suchomel, 2013) and large reference dictionaries like the OED, it may occur much more frequently within certain sociolinguistic varieties such as age groups (e.g. *buzz off* [go away] for Boomers, *delulu* [delusional] for Gen Z), geographical regions (e.g. *duck* [term of endearment] in the English midlands), gender (e.g. *mauve*, once favoured by women), register (e.g. *niggardly* might be a more erudite term than *stingy* and so much less common in everyday speech) or topic domain (e.g. *googly* in cricket or *blocking* in theatre). A term might be fairly common within a specific sociolinguistic variety, but, when communicating across varieties, that commonality is lost.

It is important to note with respect to 'Low Occurrence' that frequency is an indispensable but imperfect analytical proxy for likelihood of comprehension. For example, the N-word is far more widely understood than its attested frequencies in corpora would suggest, while *rule of law* is a moderately frequent phrase (4.2 words p/million in enTenTen21, or OED Band 5) yet many believe that it means something like a citizen's need to obey laws –

> The rule of law is not optional in the United States of America. If people want to come here, you WILL obey our laws. (US Secretary of Homeland Security Kristi Noem X 4 Feb 2025)

– rather than the supremacy of rules-based as opposed to arbitrary power, such as respecting judicial decisions and not exceeding the powers of one's branch of government (Bingham, 2011). An analytical alternative to frequency is familiarity, a hearer's subjective perception of recognition of a word, but this is also just a proxy for likelihood of comprehension. Familiarity Rating Scales (Tanaka-Ishii & Terada, 2011) actually measure three separate phenomona: familiarity with the form of the word ('I have frequently seen or heard this word'); *belief* in its meaning ('I know what it means'); and an ability to define it ('I can provide a definition for it'). Yet, as *rule of law* demonstrates, you can be familiar with the form of a term, believe you know its meaning and even believe you can define it while actually demonstrating no understanding of the term at all. Nevertheless, research has shown a clear correlation between frequency and familiarity ratings at least for spoken language (Tanaka-Ishii & Terada, 2011) and frequency is the much easier of the two to measure. For the purposes of establishing low occurrence (of an ordinary term), frequency of use can be operationalised through:

a) frequency counts in very large general corpora of a language like the TenTen Corpus family (Jakubíček et al., 2013);
b) yearly frequencies of words on Google NGram Viewer (Pettit, 2016); and
c) OED frequency bands (OED, 2025).

Corpus-based counts tend to be the most accurate in terms of contemporary usage. The Google counts, based on the enormous Google Books corpus (Juola, 2022), are restricted to usage in written books, but they are useful for tracing how a word falls in and out of favour. Finally, the OED frequency bands, which are based on the Google counts, are useful for giving meaning to the particular frequency of a word or phrase. One significant complicating factor is that we are often interested in the frequency of a particular *sense* of a word and this information can often only be gleaned from close analysis of concordances in a corpus.

4. Lexical Inferencing and Dual Processing

What happens, then, when a hearer comes across an unfamiliar LOOT? When we hear a word or phrase that is not in our mental lexicon or represented differently there, three things might happen. We might suffer a 'slip of the ear' (Bond, 1999) and hear a word that is in our lexicon rather than the one that was actually said; for example, hearing *niggard* (stingy person) as the N-word. We might perceive the word correctly, but recognise that we could not understand it without explicit external help, as in the case of technical terms like *nigrosin* (a black pigment). Thirdly, though, we might perceive the word correctly, be unable to retrieve it from our mental lexicon, but then try to infer its meaning. This is the area of lexical inferencing (Haastrup, 1991).

It is helpful at this point to draw on dual process theories of reasoning, according to which cognitive performance is the result of two types of processing, generally referred to as Type 1 and Type 2 (Evans & Stanovich, 2012; Sloman, 1996; Wason & Evans, 1974). Type 1 processes are intuitive and autonomous and so usually faster, while Type 2 processes require 'working memory' (Baddeley, 2007) and so are usually slower. Dual process theories have been particularly applied in such areas as reasoning (Evans, 2012; Evans & Over, 1996) and decision-making (Kahneman, 2013). For example, the Cognitive Reflection Test (CRT) (Frederick, 2005) tests people's ability to suppress Type 1 processing in certain non-intuitive problems. One of the CRT questions is the bat-and-ball problem:

> A bat and ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

The majority of participants incorrectly answer 10 cents because this is the rapid response suggested by our autonomous intuitive Type 1 processing. In order to get the correct answer, you need to suppress or override the intuitive response and use your working memory to actively reason through the problem reflectively (Type 2): the ball cannot cost 10 cents as then the bat will cost $1.10, making a total of $1.20; so the ball must cost 5 cents and the bat $1.05. The demonstration of such biases deriving from Type 1 processing led to the erroneous belief that Type 2 deliberative processing was necessarily 'better' than Type 1 processing. This idea has been debunked. For example, expert decision-making has been shown to depend greatly on intuitive, rather than just reflective, thinking (Gigerenzer, 2007; Gladwell, 2005). And sometimes explicit reasoning can lead to worse performance (Reber, 1993; T. D. Wilson & Schooler, 1991).

Dual process reasoning has not previously been applied to the area of lexical inferencing, but perhaps only because of a misunderstanding about how communication works. For example, Evans notes that one cannot apply Type 1 processing to *any* au-

tonomous process, such as 'those which enable visual perception and language comprehension' (Evans, 2018, p. 142). Yet, understanding language is not solely an autonomous process and Type 2 lexical reasoning processes are still available to us. For example, with technical terms such as *dysgraphia* an educated hearer might use morphological analysis to understand the term (dys- [faulty, inability] -graph- [writing] -ia [disorder]). But the ordinariness of LOOTs makes this less likely to happen. One of the most common Type 1 processes in such cases, and possibly our main means of acquiring new lexis in naturalistic contexts, is contextual inferencing (Li, 1988), or the reliance on meaning in the surrounding context to predict the meaning of the target word. In language acquisition, this is a very gradual process of intuitively piecing together the contextual evidence for the meaning of a word or phrase (rather like the accumulated contextual evidence leading to an experienced driver's ability to assess the possibility of overtaking under normal driving conditions). But even in individual cases of lexical inferencing, the hearer might successfully arrive at the correct meaning. For example, a hearer not knowing the LOOT loan word *bambino* will probably ascertain the meaning in the following utterance:

> It's the school holidays next week, so I'm taking the **bambinos** to Italy.

Indeed, very few words and phrases in our own language are explicitly learnt through looking in dictionaries, asking for the meaning, or performing structural analyses.

Another fundamental Type 1 lexical inferencing process is to rely on network relations within a 'lexical neighbourhood' in our mental lexicons. Mental lexicons are not organized like dictionaries, but rather as networks of words that are linked phonologically, semantically and grammatically. These networks form 'lexical neighbourhoods' (Vitevitch & Luce, 2016), though the place metaphor 'neighbourhood' is a little misleading since the elements are not necessarily stored in the same place in the brain. In experimental studies of speech processing, phonological neighbourhoods are defined by a one-phoneme metric, according to which the neighbour of a target word is established by the addition, deletion or substitution of a single phoneme e.g. *sit* will have amongst its neighbours *fit*, *sat*, *sin*, *sipped* and *_it* (Greenberg & Jenkins, 1964). There is a well-established consensus that similar-sounding words in a phonological neighbourhood are simultaneously activated and that there is competition for recognition among these words (Marslen-Wilson, 1987; Norris & McQueen, 2008). Psycholinguists have mostly been interested in the dense neighbourhoods of monosyllabic words and in the 'neighbourhood effects' of such density, such as 'slips of the ear', in which the listener misperceives a correctly pronounced word (Bond, 1999). But in our present case we are interested in what happens when there is no representation of the target word, phrase or word sense in the hearer's memory. Mostly we are interested in words of two or more syllables with sparser neighbourhoods, but it is what is in those neighbourhoods that is vital.

When the brain autonomously searches for the meaning of an unknown word or sense, then, it might explore the lexical neighbourhood for phonological, orthographical or sense similarity. This attempt to infer the meaning of a target word by associating it with similar words has been called an 'associating' strategy in the literature on lexical inferencing (Hu & Nassaji, 2014), but there has been no attempt to explain this strategy in terms of reasoning processes. Consider a phrase I have only recently come across: *rustle my jimmies*. This apparently emerged as an Internet meme in around 2010 and

trended until about 2015 (Jimmies, 2024). I had no idea what *jimmies* meant or what rustling them would do. I came across the term out of context and so was not able to rely on contextual inferencing.[2] Yet I still had an intuitive sense that it might mean 'annoy'. Although not explicitly aware of this, my mind probably rapidly, intuitively and autonomously searched the lexical neighbourhood of *rustle*. This produced an immediate success: *ruffle* is a phonological neighbour of *rustle*, with just a one-phoneme substitution of /f/ for /s/:

Rustle	/rʌsəl/
Ruffle	/rʌfəl/

At this point, another Type 1 linguistic process might have kicked in: collocational prediction. Given the adjective *torrential*, we intuitively predict the next word to be *rain*. Given the word *ruffle*, we are likely to intuitively think *feathers*. This can be proved empirically: by far the most frequent and strongest collocate of *ruffle* is *feathers*. In the 52-billion word enTenTen21 corpus on the Sketch Engine corpus manager (Kilgarriff et al., 2014), *feathers* (18,953 co-occurrences) occurs twice as frequently as *hair* (9,721 co-occurrences) and in terms of collocational strength (the likelihood of the node word predicting the collocate), *feathers* has a LogDice score of 10.58 to *hair*'s 6.13. The reason for *ruffle* so strongly predicting *feathers* is the common expression *ruffle one's feathers*, which means annoy. Thus there is a very strong intuitive lexicogrammatical pattern connecting the two phrases:

Verb /rʌ_əl/	+	my	+	2-syllable plural noun
Rustle	+	*my*	+	*jimmies*
Ruffle	+	*my*	+	*feathers*

This is probably not coincidental. It is likely that whoever coined the meme *rustle my jimmies* was, consciously or subconsciously, basing it on the pattern of *ruffle my feathers*. So, in this case, Type 1 intuitive reasoning allowed me to arrive at the correct answer (in a barely noticed instant). I still did not know what *jimmies* meant but, as work on formulaic language has long demonstrated (Wray, 2002), we tend to store multiword items as single units of meaning in any case (think Rice Krispies). So, in terms of communication, rather than linguistic curiosity, it was sufficient that I understood that *rustle my jimmies* meant 'annoy'.

But what if the answer does not come intuitively to the hearer through Type 1 processing? At that point they might use Type 2 reasoning by explicitly analysing the meaning of the components of the phrase (this equates with explicitly analysing the morphological components of a single word to establish its meaning). Someone from New England might have the following information in their lexical entries for *jimmy* and *jimmies*:

[2] It was suggested by an MA student on a Research Experience project as an example of a phrase designed to test AI chatbots' handling of potentially offensive words or phrases.

> *Jimmy* (noun): a short crowbar
> *Jimmies* (plural noun): 'tiny rod-shaped bits of usually chocolate-flavored candy often sprinkled on ice cream' (Merriam-Webster 2025)

Jimmies are sprinkles, or *hundreds and thousands* in British English. The cotext *rustle my* should preclude the first sense of *jimmy* in the plural (*rustle my crowbars). Sprinkles, on the other hand, are small enough to be rustled ('shake, stir, or otherwise move, causing a rustling [soft, muffled crackling] sound to be produced' OED 'rustle'). However, even knowing this, it is still not clear that the New Englander will arrive at the meaning 'annoy' without making the lexical association with *ruffle one's feathers*. Furthermore, New Englanders might also bring false etymological knowledge to the word *jimmies*: there is a persistent and widespread belief in the region that the word is racist and derives from Jim Crow, even if, as a Snopes fact check concludes (Mikkelson, 2009), 'no substantive evidence demonstrates anything denigrative of African-Americans was tied to the origin of the name.'

So, in the case of *rustle my jimmies*, Type 1 autonomous reasoning, through contextual inferencing, phonological and semantic neighbourhoods and collocational prediction, leads us to the right answer with respect to the meaning of a phrase we have not come across before. On the other hand, Type 2 deliberative reasoning using working memory can lead us astray if we lack relevant entries for key components in the phrase or if we apply false folk etymologies.

5. Offence-Adjacent LOOTs

While searching the lexical neighbourhood is often a successful Type 1 lexical inferencing process, where a LOOT is phonologically or graphically similar, or identical, to a recognised offensive term or sense in our mental lexicons, this process, like the intuitive mathematical process in the bat-and-ball problem, is likely to mislead.

Let us return to the archetypal offence-adjacent LOOT *niggardly*, as in 'I was kept on a niggardly allowance' (J. Aiken, *Blackground* 1989, in OED 2003 'niggardly' adj.). *Niggardly* derives from *nig*, a probable early Scandinavian word meaning 'parsimonious, stingy', which is related to *niggle* but completely unrelated to words deriving from post-classical Latin *nigro* ('black') such as *nigritude*, *nigrosine* and the N-word. *Niggardly* is clearly a LOOT. In terms of ordinariness, it is an adjective with a highly productive and common derivational suffix (-ly), it is clearly non-technical and it is found in everyday genres such as novels, letters and speeches. Accordingly, the speaker is more likely to presume that the hearer will understand its meaning and the hearer is more likely to presume that they should be able to infer its meaning. In terms of occurrence, while it was fairly common in the 19th century, with a frequency of 0.7 words per million (Google NGram 'niggardly'), by the end of the 20th century, when the Howard case occurred and others followed, it had fallen in frequency to 0.08 words per million, and thus the typical LOOT frequency of OED Band 3 (0.01-0.1 words p/million). Given that, in 1970, *niggardly* still had a non-LOOT frequency of 0.2 words p/million, we can infer that, by 1999, the date of the Howard case, it was probably more widely known amongst older generations (who were accused of using it) and less by younger generations (who made the complaints). Despite (or perhaps because of) the ongoing dispute over the word, it has remained at around 0.08 words p/million ever since.

Niggardly, though, also has both a phonological and orthographic neighbour that is highly offensive: the N-word. By the one-phoneme metric of neighbourliness (Greenberg & Jenkins, 1964), it is *niggard* rather than *niggardly* that is the phonological neighbour of the N-word (/nɪgə/ (British) or /nɪgər/ (US)), through addition of the phoneme /d/ (/nɪgəd/ (British) or /nɪgərd/(US)). However, the derivational suffix -ly is highly productive so that, even if there is, according to the OED, no adjectival form for the N-word (*niggerly* is only given as an alternative spelling of *niggardly*), it is very easy for us to generate this adjective in our minds. Most psycholinguistic work on phonological neighbourhoods has been based on experiments with monosyllabic CVC (Consonant-Vowel-Consonant) words (e.g. cat/bat/rat etc.) so it does not take into account the productivity of morphological derivation (i.e. our capacity to generate derived forms even when these do not currently exist in the language). Orthographically, again if we take *niggard* as the true neighbour of the N-word, we do not quite have a next-door neighbour at one grapheme distance, but we have a neighbour two graphemes down, with a substitution of [e] for [a] and the deletion of [d]: *niggard* -> *nigger_*. Furthermore, this naughty neighbour is considerably more common and familiar to people than the LOOT *niggardly*. The OED gives the frequency of the N-word at 2.5 words per million, putting it in its frequency band 5. However, OED frequencies are based on the Google Books database and thus only written discourse that has been accepted for publication. Given the extremely well-known controversy over the term, the reluctance of publishers to publish the word, and the fact that taboo terms are always far more common in speech than writing (*fuck* was common in everyday speech long before it became acceptable to use it on late-night TV or publish it in some contexts), any published frequency figures almost certainly grossly underestimate people's familiarity with the term.

6. Offence Presumption

So we have an unfamiliar LOOT that is a phonological neighbour of an extremely familiar and highly offensive slur. If we are familiar with the LOOT, this is not a problem: as we know the word, there is no need for our language processing to search the lexical neighbourhood, just as the offensive term *cunt* is not usually brought to mind when producing the word *country*. But if we are not familiar with the LOOT, then, as with the bat-and-ball problem, a rapid autonomous Type 1 process (in this case, our searching of the phonological neighbourhood in our mental lexicons) is likely to lead to an erroneous result: that the word *niggardly* is highly offensive. Furthermore, as with the bat-and-ball problem, unless something triggers conscious Type 2 reflection, we will be liable to be confident that our answer is correct. In other words, we shall **presume** that it is correct unless evidence leads us to overturn that presumption (Walton, 2014). In the former case, that evidence might be the framing of the bat-and-ball question as a 'problem', which leads around one-third of undergraduate participants to stop and reflect on their initial intuitive response. In the latter case, it might be the context that enables many hearers to stop and reflect: is the speaker really likely to have used an offensive term when describing the budget in a financial speech? However, at least some hearers will simply presume that the unfamiliar LOOT is related to its more common offensive neighbour.

There is an even more hazardous state of affairs when the LOOT is not phonologically *similar* to a lexical neighbour but is *identical* in form to the offensive term. This

is the case when the same lexeme has both non-offensive and offensive senses but the hearer only has the offensive sense in the lexical entry in their lexicon. Take a diner sitting in a restaurant looking at the menu and expressing his disgust to his fellow diners:

> 'I hate Welsh faggots!'

If the diners are from Wales, they will probably have at least two separate entries for the noun faggot in their lexicons:

> Faggot: gay man (slur)
> Faggots: meatballs made from minced off-cuts and offal

Depending on age and other factors, they might also have some of the other eighteen senses listed by the OED for the noun, such as:

- a bundle of sticks for firewood
- a small bunch of herbs for seasoning a dish
- a derogatory name for a woman (sometimes man) e.g. *old faggot*
- a naughty or mischievous child e.g. *little faggot* (OED, *faggot* n.)

In the context of dining, though, they should automatically and intuitively select the meaning referring to the Welsh dish, which is not always deeply appreciated:

> 'The contorted faces had nothing to do with the indigestibility of the school dinner faggots.' (*Western* Mail 1 March 2004) (cited in OED, *faggot* n. sense I.5)

However, imagine a gay American visitor overhearing the utterance from another table. Like the Merriam-Webster online dictionary in its entry for *faggot* (*Merriam-Webster.com*), they will probably only have the slur in their lexical entry. In the absence of alternative senses for *faggot* in their lexicon, Type 1 processing will rush to the conclusion that the speaker is either racist or homophobic or both. The hearer will simply presume that they have heard a slur rather than a neutral term. That presumption could only be overturned by Type 2 reflection on evidence relating to the context and cotext (perhaps that the company seem polite and well-mannered and that 'faggots' are on the menu) that leads to the realisation that the speaker had a different referent in mind. As with all intuitive reasoning, then, things can go wrong with Type 1 lexical inferencing, and this can lead to serious misconstruals that result in a breakdown in interpersonal relations. It is where Type 1 reasoning goes wrong with respect to lexical inferencing that we end up with offence-presumptive LOOTs.

There is an obvious objection that needs to be raised at this point. Thanks mainly to prominent cases in the US at the turn of the century, *niggardly* has, as we have seen, become controversial and, at least in the US, might be going the way of other neutral terms that have become tabooified. As one columnist put it when the controversies over *niggardly* were at their height: 'Like it or not, the word is now radioactive; having defended it, no one can now use it — especially in racially mixed company — without raising the question of motives, which, however, few will dare voice' (Poniewozick, 1999). This may well be true, but two points need to be kept in mind. Firstly, it is not clear how far high-brow discussions of words like *niggardly* permeate the population at large. Secondly, it is quite possible that it is precisely the offence-presumptive nature of the LOOT that led to the initiation of the tabooification process in the first place. There is no suggestion in the case facts of the Howard case or other early *niggardly* cases that the claimants were objecting to use of an established taboo word. It was quite clearly the

perceived link to the N-word that was at issue. Indeed, many prominent Black figures at the time ridiculed the mayor for firing Howard: the distinguished Black scholar of the N-word, Randall Kennedy, described the incident as an 'infamous round of wrongheaded protest' (Kennedy, 2002, p. 96), while the then chair of the NAACP quipped that 'the Mayor has been niggardly in his judgment on this issue' (cited in Kennedy, 2002, p. 96). In other words, in the case of offence-presumptive terms like *niggardly*, it is precisely because they presume offence that they can eventually be perceived by the community at large as offensive.

Another point to bear in mind is the frequency of commonly-believed false etymologies, particularly with regard to putatively racist terms. We saw the case of *jimmies*. Another particularly notable case is that of *nitty-gritty*, as in 'getting down to the *nitty-gritty*, or the important details', which Allan and Burridge twenty years ago thought might have suffered the kiss of death. *Nitty-gritty*, unlike *niggardly*, is not an offence-presumptive LOOT. For a start, the N-word is a fairly distant phonological and orthographical neighbour: both words start with /n/[n] and one of the syllables in both words starts with /g/[g], but *nitty-gritty* has nine phonemes (/nɪtiː-grɪtiː/), of which only three (/nɪ__g/) are shared, and four syllables (nɪ-tiː-grɪ-tiː) of which only one is shared (nɪ). Orthographically, *nitty-gritty* has eleven graphemes of which only four are shared with the N-word ([ni___gr___]). Indeed, it is not even a LOOT. It is an ordinary term: as a rhyming compound, it is extremely unlikely to be perceived as technical and so the listener is likely to presume that they can infer its meaning. But, in terms of occurrence, it does not fall in the typical LOOT band of 0.01 to 0.1 words per million. In the enTenTen21 corpus, *nitty-gritty* occurs at a frequency of 0.28 words p/million, making it a clear Band 4 word, within which 'most words remain recognizable to English speakers and are likely to be used unproblematically in fiction or journalism' (OED Frequency). There is no indication that its use is restricted to a certain demographic. With the exception of being an ordinary term, then, *nitty-gritty* does not appear to share the features of an offence-presumptive term. That is because its claim to offensiveness does not derive from the lexical neighbourhood and a low occurrence but from a false etymology, that it was originally used to refer to the detritus at the bottom of slave ships. While the word is, according to the OED, 'of uncertain origin', its first attestation is dated to 1940 and there is no evidence that it is related to the slave trade. So while offence-presumptive terms like *niggardly* (at least originally) are mistaken as offensive through faulty Type 1 intuitive reasoning, false-etymology terms like *nitty-gritty* are mistaken as offensive through faulty Type 2 deliberative reasoning: someone somewhere must have worked out the false etymology even if the claim might subsequently be taken at face value.

7. Analyzing Offence-Presumptive LOOTs

At the core of offence-presumptive LOOTs, then, is the fact that they share, or are perceived to share, a phonological or sense 'neighbourhood' with an offensive term in our mental lexicons. Not being familiar with the target word or sense, the hearer defaults to an offensive phonological or sense neighbour. Accordingly, the lexical item is simply presumed by the hearer to be offensive because they are not aware of a non-offensive alternative. And it is only if contextual or cotextual evidence triggers Type 2 reflection in working memory that the hearer will overcome that presumption. In this section I discuss the three key characteristics of offence-presumptive LOOTs and explore some of

the ways that these can be operationalised in analysis. The first two features (ordinary terms and low occurrence) derive from the fact that they are LOOTs. The third feature (having an offensive neighbour) is fundamental to the offensive-presumptive nature of these terms.

7.1. Ordinary Term

The term must be ordinary enough for the hearer, despite being unfamiliar with it, to presume that they can infer its meaning, at least to the extent that it is offensive. Hearers would not generally presume to understand technical terms such as *nigrosin* (a black pigment) or *shagbark* (a tree) and so are unlikely to presume they are offensive. *Nigre* is phonologically and graphically very close to the N-word but it is not offense-presumptive since it is only used in the technical context of soap-making (to refer to the dark-coloured solution of soap and impurities 'which settles out from the pure soap during the final stage of the manufacturing process' OED 'nigre'). Features of form and the discourse context must therefore preclude a technical meaning that hearers/readers should not be expected to know. We can also exclude proper nouns as candidates for offense-presumptiveness: they are represented differently from common nouns in our mental lexicons (Proverbio, Mariani, Zani, & Adorni, 2009); they lack the rich semantic networks of common nouns; and they tend to just store the reference, which makes them more difficult to retrieve (Semenza, 2009). While proper nouns like Dick (Van Dyke), Coon (Carrie, American actress), Cockburn (town in Scotland), Dildo (town in Newfoundland) and Shiitake (mushroom) might be sources of Type 2 explicit humour, proper nouns will not be offence-presumptive unless they are not recognized as proper nouns (e.g Spotted Dick – a traditional British steamed pudding).

7.2. Low Occurrence

The term must be infrequent enough to be unfamiliar to many hearers, but frequent enough to be familiar to the speaker and for the speaker to presume familiarity. We vary enormously in our knowledge of words so, at a whole language level, this means that there must be a strong likelihood that the hearer will not be familiar with or understand the term. Commonly understood words and phrases with offensive neighbours, such as *country*, *cocktail*, *bitch* (of dogs), *ride*, *come* and *do*, are not offence-presumptive because the hearer can easily access the non-offensive sense. In these cases, linguistic or contextual cues are required to activate an offensive frame (see section 10). For example, a radio commercial for the UK supermarket chain Somerfield was banned by the industry regulator Ofcom for a husband's utterance to his wife:

> 'I've got nothing against faggots, I just don't fancy them'

Here the issue was not that *faggot* would be offence-presumptive in a US context but that the advertisers were deliberately drawing the attention of the audience, through double entendre, both to the neutral meaning (Welsh meatballs) and to the offensive meaning (gay slur). Furthermore, the double entendre on *fancy* contextually reinforces the double entendre on *faggots*: *fancy* means 'want' but, in informal British English, it also means 'be physically attracted to,' thus making explicit the double entendre of *faggots*.

7.3. Offensive Neighbour

The final feature of offence-presumptive terms, and the feature that sets them apart from other LOOTs, is that they must share a phonological/graphological or sense neighbourhood with an offensive term or sense. Aside from *niggardly*, some other LOOTs with offensive phonological neighbours include several N-word neighbours – *niggle*, *snigger*,[3] and *chigger* (a mite common in the US) – *fag end* ('the last and worse part of something'), *gyp* ('mock' as in 'give someone gyp'), *anise* (the plant with aniseeds) and *spick and span* ('neat', 'smart'). LOOT sense neighbours or homonyms include *chink* ('fissure' v 'Chinese'), *fag* (UK, 'cigarette' v 'gay'), *faggot* (UK, 'meatball' v 'gay'), *ejaculate*[4] ('utter suddenly' v 'eject sperm'), *diddle* ('cheat, swindle' v 'copulate'), *mooncalf* ('daydreamer' v 'intellectually disabled'), *fuzzy-wuzzy* ('overly sentimental' v 'any dark-skinned person'), *butters my biscuit* ('pleases me' v 'excites me sexually') and *grinds my gears* ('annoys me' v 'copulates with me'). Sometimes, the non-offensive status of offense-presumptive items cannot be established clearly. This is particularly the case with words in the neighbourhood of the N-word deriving from the *nigro-* root indicating 'black'. For example, the now archaic term *nigritude* could be used in contexts that were not in the least offensive e.g.

> "Our aged friends can well remember when the smoke of London was not equal to one-tenth of last year's nigritude." *Quarterly Review* 168: 372 (1889) (cited in OED 'nigritude')

But there are also contexts where the association with Black people is clear, as in the following lines from a satirical poem by Victorian poet Thomas Hood:

> We've scrubb'd the negroes till we've nearly killed 'em,
> And finding that we cannot wash them white,
> But still their nigritude offends the sight.
> (T. Hood c.1845 in Black Job in Works (1862–3) vol. VI. p.166)

Over time, such contextual associations will tend to lead to the avoidance of the neutral term. The OED, for example, tells us that the N-word used to be used in several non-offensive contexts to refer to, for example: black caterpillars, ladybird larvae and sea cucumbers (II.8.a-c); a steam-driven capstan used on riverboats in the US (II.9.a); a device used to hold and turn logs in a sawmill (II.9.b); a dark brown colour (II.11); and a screen or mask used to deflect or conceal unwanted light in film studios (II.12). I doubt that the word is ever still used in these contexts though.

In the following two sections, I shall consider a couple of examples, based on forensic casework, of what might be seen as offence-presumptive terms: a figurative verb (*kneecap*) and a catchphrase (*never mind the quality, feel the width*). These claims were withdrawn before they got to court and so, in order to protect the reputations of both claimants and defendants, I shall only refer to these cases in the abstract. Although the linguistic reports together amount to almost 200 pages of analysis, including detailed

[3] *Niggle* and *snigger* are at the low end of Band 4 according to the OED, though I suspect they are more common in speech (the frequencies in the British National Corpus Spoken are too low to be significant). This might partly explain why they appear to be less controversial than *niggardly*. *Snigger* contains the entire N-word within it both phonologically and graphically but the fact that it begins with /s/ rather than /n/ might make it less offence-presumptive.

[4] Although *ejaculate* is not offensive in itself, if used in contexts referring to exclamation, a hearer not knowing this sense might find it highly offensive e.g. 'Stop ejaculating while I'm trying to speak!'

analysis of the specific communicational contexts in which the claimed offence took place, since it is not possible to adduce the detailed case-relevant evidence, I will simply ask the reader to accept hypothetical scenarios based on the cases. At the core of these hypothetical scenarios is a situation in which the claimant has genuinely taken serious offence on hearing a word or phrase uttered by the defendant, who in turn is completely oblivious to the possibility that they might have said anything offensive. What is important here is not the rights and wrongs of the specific cases but the potential explanatory value of the proposed category of offence-presumptive LOOTs.

8. Presumptive Racism: Figurative *Kneecap* (v.)

In the first case, imagine two experts discussing disciplinary measures with respect to a professional infraction and the defendant, considering that the claimant is being too harsh, saying something like: 'Do you want [the profession's regulatory body] **to kneecap** the poor [professional].' Imagine also that the claimant, unbeknownst to the defendant, is from Ireland and that they feel a strong association between the verb *kneecap* and the Troubles in Northern Ireland in the 1970s and 1980s. The claim, then, is that use of the verb *kneecap* by the defendant is racist.[5] Unable to discuss the details of the case, I would ask the reader to simply *assume* for the sake of argument that there are no circumstantial or contextual linguistic details that point to racism, or rather xenophobia, beyond the denotational and connotational meaning of the verb.

8.1. Overview of the Meaning of the *Verb Kneecap*

The verb *kneecap* was coined to describe the shooting of a person in the knee as a form of extrajudicial punishment by paramilitary groups and was, according to the OED, first attested in a newspaper headline in 1975:

> Man 'kneecapped' in Carrickfergus (*Daily Telegraph* 12 August 1975)

During the Troubles, the verb will certainly have developed extremely negative connotations, even if those connotations do not appear to have made it into dictionary entries for the verb. Over time, though, the meaning extended figuratively to hyperbole, as in the following description of attempted tackles in football/soccer:

> He'd be running from the halfway line with defenders trying to **kneecap** him. (*The Daily Telegraph* 22 March 2013)

And then, most commonly, to metaphor:

> This comes as Iran's crude exports have been **kneecapped** by US sanctions, which were re-imposed in November. (*S&P Global Commodity Insights*. 16 May 2019)

We can observe a clear pattern of use through the form's collocations. Figure 1 shows the top 20 collocates, scored by Mutual Information (MI) (the mutual dependency of one word on the other), for the word form 'to kneecap' in enTenTen21:

[5] Strictly speaking, it would be xenophobic rather than racist but the claim was for 'racism'.

	Word	Cooccurrences	Candidates	MI
1	Mueller	8	186,343	11.76
2	competitor	6	358,250	10.40
3	rival	8	702,303	9.85
4	Sanders	5	449,772	9.81
5	opponent	6	880,717	9.11
6	opponents	6	903,791	9.07
7	unions	5	809,102	8.97
8	Republicans	7	1,194,057	8.89
9	Trump	14	3,401,506	8.38
10	Obama	7	2,579,886	7.78

	Word	Cooccurrences	Candidates	MI
11	economy	5	4,566,229	6.47
12	efforts	5	5,259,073	6.27
13	ability	6	6,929,890	6.13
14	political	8	9,647,962	6.07
15	anyone	6	8,580,772	5.82
16	him	31	44,951,310	5.80
17	themselves	5	8,602,781	5.56
18	American	7	15,348,465	5.21
19	every	7	23,336,672	4.60
20	me	17	58,332,008	4.56

Figure 1. Top 20 Collocates of "To Kneecap" in enTenTen21 by MI Score (Tokens=761; Range=R1-3; Min. Freq.=5)

We can see that 'to kneecap' collocates strongly with words indicating opponents (*competitor, rival, opponent(s)*) and specified (particularly US) political opponents (*Mueller, Sanders, Trump, Obama*), or opposing parties (*unions, Republicans*), all of whom have been metaphorically rather than physically kneecapped. The *economy, efforts* and *ability* have also been metaphorically debilitated. There are no Troubles-related collocations (though these do exist for the more historical form *kneecapped*). 'To kneecap', therefore, primarily appears to indicate an intention to figuratively debilitate opponents and it seems to have lost its negative connotations associated with the paramilitary groups that first developed the practice.

8.2. Ordinary Term?

Kneecap is clearly an ordinary term. The noun on which it is based is almost universally understood, unlike its medical synonym *patella*. The verb was coined via the very common derivational process of converting an ordinary noun to a verb (e.g. *to fax, to facetime, to studioproduce*). Although the verb initially referred to a very specific form of extrajudicial punishment, it was never used technically and was first attested in the media, probably reporting spoken usage within the community. It later extended in meaning both in the scope of its literal coverage (any deliberate wounding of the leg in any context) but also in its extension to figurative uses (hyperbole and metaphor). The verb appears to be found predominantly in the media, though a lack of sizable spoken corpora mean that it is not clear how extensive its use is in spoken language. In any case, the verb is ordinary enough for the speaker to presume that it does not need explaining and for the hearer to presume that they would be able to know or infer its meaning.

8.3. Low Occurrence?

Establishing the frequency of the verb *kneecap* is difficult because neither the OED nor enTenTen20[6] distinguishes between noun and verb uses of *kneecap*. Manual analysis of citations from searches of various corpora reveals, though, that the verb *kneecap* accounts for approximately one-fifth or one-sixth of the total occurrences of *kneecap*. If we divide the frequency of *kneecap* by one-fifth (the estimated proportion of the verb

[6] EnTenTen20 rather than 21 was used for this analysis since that was the latest TenTen corpus available at the time of the analysis.

kneecap to overall *kneecap* in enTenTen20), we arrive at a frequency of 0.068 per million, which places it firmly within the OED's Band 3 and at a similar frequency to *niggardly*. However, we then need to separate the literal and figurative senses. In a random sample of the occurrences of the verb *kneecap* in enTenTen20, 83% were found to be figurative (hyperbole or metaphor) and 75% specifically metaphorical. This might suggest that the figurative sense (the one that is supposedly offence-presumptive) should be more familiar to a claimant because the figurative sense has become the 'dominant' meaning of the verb while the literal sense is now 'subordinate' (Simpson, 1981). However, in the 1993 subcorpus of the English Broadsheet Newspapers 1993-2013 corpus (also on SketchEngine), 75% of the uses refer to kneecapping during the Troubles, thus suggesting that this was the dominant meaning at the time. If a claimant acquired their lexical entry for *kneecap* in Ireland in the 1980s or 1990s, then it is likely that the literal Troubles-related sense is still the one familiar to them, while the low frequency of the verb as a whole means that it is quite possible that they are unfamiliar with the now much more common figurative sense. So, although figurative *kneecap* is now far more frequent overall than literal *kneecap*, an Irishman of a certain age is perhaps going to be more familiar with the literal sense than the figurative one.

8.4. Offensive Neighbour?

Clearly we have a *faggot*-type situation here where the hearer may simply not have one of the senses in their mental lexicon and so defaults to the other sense. With *kneecap*, the same lexeme has extended in meaning from a literal sense (shoot in the knees) originally associated with the Troubles in Northern Ireland to a metaphorical sense (bring to its knees) unrelated in any way with the Troubles. The association of *kneecap* with Northern Ireland is clearly still alive within that nation, as is evident from the name of one of their leading (if controversial) bands: Kneecap. Furthermore, the only three references to literal *kneecap* associated with the Troubles in the large Nexis database of current news sources for the period 2020-22 are from Northern Ireland newspapers. Outside Northern Ireland, though, the corpus evidence suggests that the connection between the verb *kneecap* as a whole and the Troubles has been entirely lost.

The question, though, is whether the sense neighbour of the now-dominant figurative use, i.e. literal *kneecap*, can be considered offensive. On the one hand, rather than becoming taboo after the Troubles, the literal sense has become extended to encompass a wider range of crimes than simply paramilitary shooting in the leg. Furthermore, no dictionary marks up *kneecap* as derogatory either in its literal or figurative usage and there is no evidence of the verb *kneecap* being discussed as potentially derogatory in itself. Where terms are known to be offensive, such as the noun use of *cripple*, or believed by some to be offensive, it is generally not difficult to find discussion of this on the Web. On the other hand, for someone who lived through the Troubles, it is not difficult to see how the word might be considered offensive within their particular community. Moreover, that perception of offensiveness might make them more resistant to accepting a non-offensive figurative use of the verb. To the potential objection that the claimant should have realised that the verb must have been being used in a figurative sense, one can respond that almost any verb can be used in a figurative sense but, unless the verb is commonly used figuratively in similar contexts, it can be hazardous to do so. For example, consider if the defendant, instead of using figurative *kneecap*, had

tried to use *shoot* figuratively: 'Do you want them to *shoot* him?' This would have been inexplicable to the hearer because we would not use *shoot* figuratively in this context. So a claimant, lacking a non-offensive figurative sense in their mental lexicon, might apply the same negative connotations of the literal Troubles-adjacent sense of *kneecap* to a context where the more neutral metaphorical sense of *kneecap* is intended. Once again, Type 1 autonomous intuitive reasoning will lead the hearer to default to their perceived offensive literal sense rather than reflect more deliberatively on the more likely pragmatic meaning in the specific context.

In conclusion, the metaphorical sense of the verb *kneecap* is offence-presumptive because it is a Low-Occurrence Ordinary Term that shares a literal sense neighbour which, for a claimant who acquired the word in Ireland in the 1970s or 1980s, has strongly negative offensive connotations associated with paramilitary terrorists in Northern Ireland. Contextually, the defendant probably felt he was using a standardly available verbal resource for conveying his opposition to a proposed harsh form of punishment, while the claimant probably felt that his identity as an Irishman was being invoked offensively through use of an 'offensive' term indexing paramilitary brutality.

9. Presumptive Sexual Harassment: *Never mind the quality, feel the width*

The second case involves presumed sexual harassment in a business meeting. Imagine a male defendant, on seeing the agenda move from a section titled 'Quality' to a section titled 'Quantity', quipping 'never mind the quality, feel the width' and a female attendee being deeply offended by the utterance of that phrase.

9.1. Overview of the Meaning of the Catchphrase

The catchphrase derives from the popular British 'cultural clash comedy' (Vice 2021: 188) *Never Mind the Quality, Feel the Width* (1967-71) about a Jewish and Catholic tailor who go into business together. According to *Brewer's Dictionary of Phrase and Fable* (Ayto & Crofton, 2011), the title 'may have been a deliberate inversion of a cloth trade saying: 'Never mind the width, feel the quality", though I was unable to find any evidence of this supposedly original saying. On the other hand, contrasting quality with quantity using a phrase beginning 'Never mind the quality…' does precede the sitcom by many decades, as in this advice to nineteenth-century pianists from a music magazine:

> Big tone is the companion to technic. So it is big, **never mind the quality; only give us quantity**. (*The Etude*, June 1888)

Semantically, then, the phrase suggests that quantity is more important than quality. What the sitcom seems to have introduced, though, is an ironic reading so that, in contemporary pragmatic usage, the catchphrase is actually almost always suggesting that we should be prioritizing quality over quantity. This is probably why it seems to be most common in the Hansard archive of UK parliamentary debates (Hansard, 2024). For example:

> Most importantly—this point constantly escapes those who take a **never mind the quality, feel the width** approach to investment—the quality of British investment is vastly improved. (PM John Major, Autumn Statement, 23 January 1990, Commons)

> The way the Minister replied—which is obviously in her brief—was all about the numbers: **never mind the quality, feel the width**. We are talking about quality of life, integration and the chances that a person who gets status would have to thrive in the UK. (Baroness Ludford, Nationality and Borders Bill, 8 February 2022, Lords)

It also explains why it has been used in the titles of academic papers:

> Binns, A. & Potter, R. (1989) Improving the effectiveness of postgraduate supervision: **never mind the quality, feel the width**, *Journal of Geography in Higher Education* 13(2): 211-16.

The phrase is therefore used ironically, though in serious contexts, to critique the privileging of quantity over quality.

9.2. Ordinary Term?

The phrase consists in extremely common words and, unlike the technical legal formula *beyond reasonable doubt*, the words in the phrase are being used in their ordinary senses. The hearer needs to understand that the phrase is being used ironically, but this is not uncommon with catchphrases. Although it derives from a title, there is nothing in the phrase itself to indicate that it is the title of a sitcom. Also, the high productivity of the phrase indicates that it has in fact become an established catchphrase rather than just a reference to a sitcom. The examples below depend for their understanding on the reader being able to make the link with the original catchphrase, helped by the phonological (/wɪt/ > /wɪdθ/) and rhyming (/wɪdθ/ > /mɪθ/) near-neighbours:

> **Never mind the quality, feel the wit**. As Sir Noel Coward remarked in rhyme, "The plot of La Gioconda is apt to wander", a somewhat optimistic judgment: in fact, it does not even get out of the front door and on to the pavement. (*The Times* 4 May 1993)

> THE plethora of tributes and retrospectives which have attended their reunion suggests that the Velvet Underground were always as much a media phenomenon as they were a musical one. **Never mind the quality, feel the myth**. (*The Daily Telegraph* 3 Jun 1993)

There is no reason, then, for a hearer to presume that they have not understood the phrase. It uses everyday words, is clearly non-technical, is an obvious catchphrase and is highly productive.

9.3. Low Occurrence?

The occurrence patterns make it highly likely both that a hearer would not be familiar with the phrase and that a speaker would expect familiarity. In terms of unfamiliarity, the phrase does not appear in US contexts (there are no occurrences in the Corpus of Contemporary American English (COCA) or in the US domain of enTenTen20) so an American hearer is unlikely to know the phrase. Furthermore, where the phrase does occur in corpora, it ranges in frequency from 0.02 words p/million (enTenTen20 UK Domain), through 0.05 words p/million (British National Corpus; English Broadsheet Newspapers 1993-2021) to a range from 0.05 to 0.1 words p/million in UK Parliamentary debates (Hansard, 2024). In other words, the frequencies are all in the OED Band

3 LOOT frequency range (0.01 to 0.1 words p/million). Turning to expected familiarity, the productivity of the phrase when used in the media and on the web suggests a perception by writers that readers will be able to understand the root catchphrase that is being manipulated. Moreover, there is some evidence of age restriction. While the phrase seems to be most common in the Hansard records of UK parliamentary debates (Hansard, 2024), most of the speakers recorded as using the phrase are now in their 70s or 80s. This means they would have been in their teens or twenties when the sitcom aired in the late 1960s. There is no indication, on the other hand, of gender restriction: many of the users of the phrase in the Lords are baronesses. Accordingly, it is likely that speakers over a certain age would consider the phrase well understood.

9.4. Offensive Neighbour?

Extensive lexicographical and corpus searches revealed no offensive sense (even an emerging one) of the phrase. However, there is a recognized sexually allusive phrase in the lexical neighbourhood. Here it is not a phonological neighbour but a pragmatic one involving slightly complex semantic and grammatical relations:

> 'It's not the size that matters, it's what you do with it'

The first half matches semantically with 'never mind the width' and the second half (less successfully) with 'feel the quality'. Then *quality* and *width* need to be inverted:

> It's not the size that matters -> never mind the ~~width~~ -> quality
> It's what you do with it -> feel the ~~quality~~ -> width

Remember that this would all be occurring through Type 1 autonomous intuitive reasoning. It is also possible that an erroneous Type 2 reasoning process might come into play: it has been suggested to me that the mere mention of the sitcom is taboo because the show was supposedly so utterly sexist. However, if the show is now perceived as controversial, it is because of the stereotyping of religious identity (though it was praised by some religious organizations at the time for promoting religious tolerance), rather than excessive sexism. This was no *Benny Hill* (Gibbs, 2023).

In conclusion, although there is no linguistic evidence that the catchphrase *never mind the quality, feel the width* is offensive, its ordinary term status, low occurrence, distribution bias towards the elderly and exclusion of the US, and its sharing not a close but a perceivable neighbour with an attested offensive phrase (at least in the context of a business meeting) can perhaps explain how a claimant might have perceived the phrase to be offensive.

10. Priming as Activating an Offensive Frame

At this point we can tackle the last step in the argument. One question that arises in such forensic cases as *kneecap* and n*ever mind the quality, feel the width* (or the early *niggardly* cases in the US or *nitty-gritty* in the UK) is why they are not dismissed as simple misconstruals by managers, HR personnel and the police. One possibility is the legal stipulation that it is the claimant who decides whether conduct is 'unwanted' and the understandable fear, then, of challenging the claimant. Yet, in my experience, even informants completely unconnected with the legal cases tend to see offensiveness in these terms when that is raised as a possibility. Instead, it is possible that those

dealing with such cases are metalinguistically **primed** to activate an offensive framing for the linguistic context. With some genres such as sitcoms, comedy shows and adverts, it is the expectations of the genre that prime audiences to be extremely attentive to ambiguous meanings and double entendres, so that even mildly suggestive material can provoke strong humorous reactions. But explicitly raising awareness of a potential offensive meaning can also prime interlocutors to activate that meaning. When I have informally asked informants whether the phrase *never mind the quality, feel the width* might convey sexual innuendo, if they do not know the phrase they inevitably reply in a somewhat tentative affirmative ('I can see how it could'). They generally cannot put their finger on why that might be the case, just as someone hearing the phrase *rustle my jimmies* might think 'annoy' while not knowing quite why, since this is an instinctive Type 1 process. The more they dwell on it, though, the more likely they are to think that the phrase is offensive. Yet MPs and Lords would not use a phrase in Parliament that they believed might convey sexual innuendo. It seems clear that these informants, like the audience in a pantomime, are being primed to see innuendo.

An example of how one can become hypersensitized to sexual innuendo, and the dangers of this for misconstrual, came in my analysis of the sample concordance lines of *it is not the size that matters…*. With my attention heightened to sexual innuendo, I came across the following example from an interview (all items that could potentially activate a sexual innuendo frame are in italics):

> HME: Why is it important for providers to be part of AAHomecare?
> Steedley: The association is *impotent* if it doesn't have *members*. When *staff* is on Capitol Hill, we have to speak for our *membership*, and our *membership* has to be reflective of the industry. We need *members of all sizes*. ***It's not the size that matters***, but the voice and *passion*. We need *members* to get legislation passed, to get meetings with CMS, to get regulatory *relief*. Otherwise, the program *laid on top of us* will continue to *grow* without any *relief* in sight. (HMEnews.com)

After analyzing many other examples of the phrase that obviously conveyed sexual innuendo, it seemed too much of a coincidence that this passage could be so packed with words and phrases with what might be called Sexual Innuendo Potential without there being any attempt actually to convey sexual innuendo. Although the overall content did not seem to suggest innuendo, I even wondered whether this was from a satirical site. Yet the website HME News (https://www.hmenews.com/) is a very sober site for 'Business News for Home Medical Equipment Providers.' There is nothing in the interview or on the site itself that would suggest that there was any conscious effort to convey sexual innuendo. Indeed, I subsequently realised that, like some examples of *never mind the quality* found on its own, here *it's not the size that matters* almost certainly is not meant to recall the sexually connotative phrase (which requires some variant of 'do with it') at all. Returning to this example now, after a couple of years, it seems extraordinary that I could have seen sexual innuendo in it, but this demonstrates the power of priming.

11. Conclusion and Implications for Judging Linguistic Offence

In this article, I have argued that linguistic offence through offence-presumptive Low Occurrence Ordinary Terms (LOOTs) comes about when a hearer is not familiar with a word or phrase and intuitively uses Type 1 lexical inferencing strategies to guess its meaning. In doing so, they draw on an offensive neighbour (lexeme or sense) in their

mental lexicons and transfer the offensiveness of that neighbour to the neutral target word. Accordingly, they are subjectively genuinely offended but objectively the offence is mistaken, just as it would be if the claimant misheard what was actually said. It should be stressed that this is a theoretical argument rather than an empirical claim. While I have drawn on empirical evidence wherever I could find it, there are very clear empirical gaps that I have had to plug in purely theoretical terms. For example, while there is a great deal of psycholinguistic empirical work on lexical neighbourhoods in relation to the question of lexical access to items already in our lexicons, this work has not been extended to cases where the target items are missing from the lexicon and so have to be inferred. Similarly, dual processing theories have not as yet been applied empirically to lexical inferencing, which would seem to be an important avenue to research. The notion of LOOTs is more empirically robust and has been developed over years of casework. However, the precise effects of these LOOTs on offence presumptiveness is currently purely theoretical. It also should be acknowledged that for radical eradicationists like Stojnić and Lepore (2025), it would be sufficient for an offence-presumptive term to be articulated, even in the context of this article, for it to cause offence. However, as with N-word use eradicationism, this appears to have remained a very minority view with respect to the public at large, who tend to be contextualists when it comes to their views of taboo language (Jay & Janschewitz, 2008).

If the argument here is sound, a number of implications follow for forensic lexical analysis in general, for the particular case of judging linguistic offence, or public claims that a word or phrase has caused offence rising to the level of remedy-required harm (Heffer, *in prep*) and for the ethics of interpersonal communication more broadly (Heffer, 2020).

With regard to forensic lexical analysis in general, the lexical category introduced here of the LOOT (Low Occurrence Ordinary Term) appears to be central in cases where the meaning of words and phrases is in dispute. It occurs in trademark cases, where the issue is often whether the average consumer is likely to consider the trademarked term as an ordinary word or phrase in the language or only as a commercial brand, as in *Bambino* (Heffer, 2008). And it occurs, as we have seen, in cases of disputed offensiveness. Empirical work is required, though, to establish precisely what happens in terms of lexical inferencing when a hearer (or reader) encounters an unfamiliar LOOT. It would also be interesting to drill down into specific sociolinguistic variables to test my hypothesis that those who use LOOTs belong to specific discourse communities in which the frequency of the term is much higher than in the general language community (such as elderly UK politicians and policymakers for *never mind the quality…*) but also that hearers who are likely to be offended by these LOOTs also belong to specific communities (such as those who lived through the Troubles for *kneecap* or eradicationists for LOOTs adjacent to the N-word).

With respect to disputes over linguistic offence, claims still need to be judged against the 'linguistic facts' of how a word or phrase is currently being used across society, as evidenced in large corpora of the language. Dictionary entries and particularly etymologies are not sufficient for gaining a clear picture of those facts. We need to search large general reference corpora to gain some idea of how a word or phrase is actually being used. We also need to put our own ideological beliefs about taboo words to one side: a certain word might be considered taboo in a given highly restricted domain,

but it would not be reasonable for someone within that domain to expect others outside it to conform to their beliefs about the word. For example, while London's Metropolitan Police might have bought into the false etymology of nitty-gritty in the 1990s, it was unreasonable to severely and publicly admonish a UK government minister for using the word in a speech he gave to a police conference in 2002 (Hopkins, 2002). A quarter of a century later, the question is still far from settled: while Sky banned its sports journalists from using *nitty-gritty* in 2020, a year later the BBC rejected a complaint about the journalist Laura Kuenssberg using the word (Moore, 2021). *Nitty-gritty*, like *niggardly*, might eventually become taboo across society, at which point this will emerge from the corpus evidence and their etymology will no longer be considered a sufficient defence, but until such point one needs to be conservative in judging the taboo status of a word that is in dispute.

While it is right that legislation on workplace harassment in both the UK and the US makes clear that 'unwanted conduct' 'means "unwanted by the worker" and should be considered from the worker's subjective point of view' (EHRC, 2020), the law also requires, once a case reaches a workplace tribunal, that the claim be deemed 'reasonable' in the circumstances. Yet an enormous amount of time and money can be expended, and lives torn apart, before the reasonableness of the claim is taken into account. It would be much better if, as with patients presenting to A&E, a case could be 'triaged' by HR personnel using a heuristic like the RESPECT framework (Heffer, In prep) to filter out cases where a claim does not correspond to the linguistic facts and has no circumstantial elements to support it. Such triaging would require systematic reflection, via the heuristic, to overcome potential metalinguistic priming. It should be stressed that, in proposing such triaging, I am concerned, as with all researchers in this area, with helping reduce the prevalence of language-based workplace harassment. A study in *The Lancet* found that 1 in 10 employees reported experience with workplace bullying or harassment in the previous year (Bunce et al., 2024), and while language-based harassment only constitutes a part of that total, it still corresponds to too many cases. Seriously tackling harassment in the workplace, though, also means dealing with those cases where current procedures misfire. Like vexatious claims, mistaken claims of linguistic offence do not just seriously harm the unjustly accused, but they also undermine confidence in procedures for identifying and judging the claims and, if common, can lead to a backlash in society.

Finally, with regard to the ethics of interpersonal communication more broadly (Heffer, 2020), since communication can be hazardous and mistakes are going to be made, we need both charity and apology. Applying the principle of charity in interpretation (N. L. Wilson, 1959) means considering the most favourable contextual interpretation of a speaker's utterance before escalating a claim of offence. If the claim of offence is reasonable, it will withstand such steelmanning. Apologizing, and accepting the apology, on the other hand, is the standard way of repairing breakdowns in interpersonal communication. Even when the speaker knows that what they said is objectively not offensive, if it is clear that the hearer is subjectively offended, then an apology is warranted. At the same time, neither the hearer nor third parties should take that apology as an admission of guilt that then warrants escalation of the claim. The more one can avoid interpersonal harm or the escalation of claims of linguistic offence

based on offence-presumptive terms and false etymologies, the more one can focus on those many justifiable cases of grievance that warrant ethical or legal remedy.

References

Allan, K., & Burridge, K. (2006). *Forbidden Words: Taboo and the Censoring of Language*. Cambridge: Cambridge University Press.

Ayto, J., & Crofton, I. (Eds.). (2011). *Brewer's Dictionary of Modern Phrase & Fable*. Online: Chambers Harrap.

Baddeley, A. (2007). *Working Memory, Thought, and Action*. Oxford: Oxford University Press.

Billig, M. (2013). *Learn to Write Badly: How to Succeed in the Social Sciences*. Cambridge: Cambridge University Press.

Bingham, T. (2011). *The Rule of Law*. London: Penguin.

Bond, Z. S. (1999). *Slips of the Ear: Errors in the Perception of Casual Conversation*. New York: Academic.

Bunce, A., Hashemi, L., Clark, C., Myers, C.-A., Stansfeld, S., & McManus, S. (2024). Prevalence and nature of workplace bullying and harassment and associations with mental health conditions in England: A cross-sectional probability sample survey. *The Lancet*, *24: 1147*.

Cameron, D. (1995). *Verbal Hygiene*. London: Routledge.

Coulthard, M. (2004). Author identification, idiolect, and linguistic uniqueness. *Applied Linguistics*, *25*(4), 431–447.

Derbyshire, J. (2002, September 17). Niggling doubts. *National Review Online*.

de Ruiter, J.-P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*(3), 515–535.

Dickerson, D. (1999, February 5). The last plantation: The 'niggardly' scandal should teach whites to watch their language and blacks to toughen up. *Salon*.

Dinn, W. M., & Harris, C. L. (2000). Neurocognitive function in antisocial personality disorder. *Psychiatry Research*, *97*(2-3), 173–190.

Dynel, M. (2018). *Irony, Deception and Humour: Seeking the Truth about Overt and Covert Untruthfulness*. Berlin: Mouton de Gruyter.

EHRC. (2020). *Sexual Harassment and Harassment at Work* [Equality and Human Rights Commission].

Evans, J. (2012). Dual-process theories of deductive reasoning: Facts and fallacies. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 115–133). Oxford: Oxford University Press.

Evans, J. (2018). Dual Process Theory: Perspectives and Problems. In W. De Neys (Ed.), *Dual Process Theory 2.0* (pp. 137–155). London: Routledge.

Evans, J., & Over, D. E. (1996). *Rationality and Reasoning*. Hove, U.K: Psychology Press.

Evans, J., & Stanovich, K. E. (2012). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Gibbs, J. (2023). A modern version of restoration comedy? Double entendre, objectification, fearful men and rakes manqué in the television work of Benny Hill. *Critical Studies in Television*, *19*(4), 468–485.

Gigerenzer, G. (2007). *Gut Feelings: The Intelligence of the Unconscious.* New York: Penguin Books.

Gladwell, M. (2005). *Blink: The Power of Thinking without Thinking.* New York (NY): Little, Brown and Co.

Grant, T. (2013). TXT 4N6: Method, consistency, and distinctiveness in the analysis of SMS text messages. *Journal of Law and Policy*, *58*(1), 467–494.

Greenberg, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, *20*(2), 157–177.

Grice, P. (1975). Logic and conversation. In *Syntax and Semantics 3: Speech Acts* (pp. 41–58). New York: Academic Press.

Gumperz, J. J. (1993). Types of linguistic communities. *Anthropological Linguistics*, *35*(1/4), 130–142.

Haastrup, K. (1991). *Lexical Inferencing Procedures or Talking about Words.* Tübingen: Gunter Narr.

Hansard. (2024). *Hansard Online Archive.* Retrieved from https://hansard.parliament.uk.

Hansen, S. J., McMahon, K. L., & De Zubicaray, G. I. (2019). The neurobiology of taboo language processing: fMRI evidence during spoken word production. *Social Cognitive and Affective Neuroscience*, *14*(3), 271–279.

Happé, F. G. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbolic Activity*, *10*(4), 275–295.

Heffer, C. (2008). *The Perceived Meaning in the UK of the words "Bambino Mio".* Bambino Mio Ltd and Cazitex.

Heffer, C. (2020). *All Bullshit and Lies? Insincerity, Irresponsibility and the Judgement of Untruthfulness.* New York: Oxford University Press.

Heffer, C. (In prep). *Linguistic Offence: The RESPECT Framework for Evaluating Disputed Offensive Language.* Cambridge: Cambridge University Press.

Hopkins, N. (2002, May 15). Why nitty gritty has been ruled a no-no in the police lexicon. *The Guardian.*

Hu, H.-c. M., & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System*, *45*, 27–38.

Hudson, R. (1996). *Sociolinguistics.* Cambridge: Cambridge University Press.

Jackendoff, R. S. (2002). *Foundations of Language: Brain, Meaning, Grammar, and Evolution.* Oxford: Oxford University Press.

Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen corpus family. In *7th International Corpus Linguistics Conference CL* (pp. 125–127).

Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research. Language, Behaviour, Culture*, *4*, 267-288.

Jimmies. (2024). *That really rustled my jimmies.* Retrieved from https://knowyourmeme.com/memes/that-really-rustled-my-jimmies.

Juola, P. (2022). Google Books Ngrams. In L. A. Schintler & C. L. McNeely (Eds.), *Encyclopedia of Big Data.* Berlin: Springer.

Kahneman, D. (2013). *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux.

Kennedy, R. (2002). *Nigger: The Strange Career of a Troublesome Word.* New York:

Pantheon Books.
Kies, D. (1985). Some stylistic features of business and technical writing: the functions of passive voice, nominalization, and agency. *Journal of Technical Writing and Communication*, *15*(4), 299–308.
Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., … Suchomel, V. (2014). The Sketch Engine. *Lexicography*, *1*(1), 7–36.
Kurtić, E., Brown, G. J., & Wells, B. (2013). Resources for turn competition in overlapping talk. *Speech Communication*, *55*(5), 721–743.
Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.
Li, X. (1988). Effects of contextual cues on inferring and remembering meanings of new words. *Applied Linguistics*, *9*(4), 402–413.
Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102.
McDonald, L. (2020). Your word against mine: The power of uptake. *Synthese*, *199*(1-2), 3505–3526.
McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14*(6), 648–652.
Mikkelson, B. (2009, September 25). Were jimmies named after Jim Crow. *Snopes*.
Moore, M. (2021, January 23). BBC rejects race complaint about use of 'nitty-gritty'. *The Times*.
Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395.
OED. (2025). *Frequency.* Oxford: Oxford University Press. Retrieved from https://www.oed.com/information/understanding-entries/frequency/?tl=true
Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology*, *20*(2), 139–156.
Pettit, M. (2016). Historical time in the age of big data: Cultural psychology, historical change, and the Google Books Ngram Viewer. *History of Psychology*, *19*(2), 141–153.
Pinker, S. (2007). *The Stuff of Thought: Language as a Window into Human Nature*. London: Penguin.
Poniewozick, J. (1999, February 2). The little N-word. *Salon*.
Proverbio, A. M., Mariani, S., Zani, A., & Adorni, R. (2009). How are 'Barack Obama' and 'President elect' differentially stored in the brain? An ERP investigation on the processing of proper and common noun pairs. *PLoS ONE*, *4*(9), e7126.
Reber, A. S. (1993). *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford: Oxford University Press.
Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and Thought* (2nd ed., pp. 164–201). Cambridge: Cambridge University Press.
Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696.
Saussure, F. d. (1983). *Course in General Linguistics*. London: Duckworth.
Semenza, C. (2009). The neuropsychology of proper names. *Mind & Language*, *24*(4),

347–369.
Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423.
Shannon, C. E., & Weaver, W. (1964). *The Mathematical Theory of Communication*. Chicago: University of Illinois.
Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior*, *20*(1), 120–136.
Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22.
Snow, T. (1999, February 3). Linguistic lynching over 'niggardly'. *Des Moines Register*.
Stojnić, U., & Lepore, E. (2025). *Inflammatory Language: Its Linguistics and Philosophy*. Oxford: Oxford University Press.
Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
Tanaka-Ishii, K., & Terada, H. (2011). Word familiarity and frequency. *Studia Linguistica*, *65*(1), 96–116.
Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, *2*(1), 75–94.
Walton, D. N. (2014). *Burden of Proof, Presumption and Argumentation*. New York (NY): Cambridge University Press.
Wason, P., & Evans, J. (1974). Dual processes in reasoning? *Cognition*, *3*(2), 141–154.
Wilson, N. L. (1959). Substances without substrata. *Review of Metaphysics*, *12*, 521–539.
Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, *60*(2), 181–192.
Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
Wray, A., & Staczek, J. J. (2025). One word or two? Psycholinguistic and sociolinguistic interpretations of meaning in a civil court case. *International Journal of Speech, Language and the Law*, *12*, 1–18.