

The metalinguistics of offence in (British) English: A corpus-based metapragmatic approach

Abstract

Offence is a central concept in impoliteness, aggression and conflict research, yet has received only passing mention in definitions of impoliteness and related concepts. Janicki (2017) argues that impoliteness and language aggression scholars are needlessly worried about definitions. We use Janicki's (2017) work as a spring-board into a discussion of definitions of impolite or taboo language, airing potential problems and suggesting that the study of metalanguage offers at least a partial solution. We report a study of the metalanguage of OFFENCE in British English, and briefly examine whether there are any differences in Australian English, using SketchEngine to interrogate data in the two-billion word Oxford English Corpus. In so doing, we tease out different uses of the term *offensive*, and show that concepts such as OFFENCE are coloured by the specific linguistic and cultural contexts in which they appear. We conclude that while corpus-based metalinguistic analyses cannot completely eliminate the problem of definitional infinite regress, they do, however, offer an empirically grounded way of defining words that allows us to move beyond the intuitions of individual researchers.

Keywords

impoliteness, offence, definition, metalanguage, metapragmatics, corpora, British English, Australian English

1. Introduction

Like all fields of social scientific inquiry, defining our object of study in analyses of aggression and conflict is invariably a complex endeavour. Yet for those particularly interested in the role of language there is an added layer of complexity. Language is not only our object of study, but the tool by which we study it. It has traditionally been assumed that we can maintain a distinction between these through the use of technical language. Indeed, as Jucker (in press) notes, "technical terms have the advantage over present-day concepts that they can be given more precise definitions and thus serve as useful labels for clearly delimited sets of phenomena" (p.12). However, scientific understandings of language very often draw from lay understandings, and so can be readily confounded with them (Scheff 2006), a point that has been repeatedly made with respect to (im)politeness research (Eelen 2001; Watts, Ide and Ehlich 1992). While we can, of course, carefully distinguish between everyday uses of terms as one of our objects of study, and the technical use of terms to define the scope and ontological terms of our study, this begs the question of what underpins definitions of the latter. While one might argue that the choice of technical label itself is a relatively trivial matter (Jucker in press: 14; cf. Watts 2005: xv), our definitions of technical concepts are most definitely not. Language is used to structure our social reality (Berger and Luckmann 1966). When we categorise and describe social behaviour as instances of impoliteness, aggression and conflict, we are making ontological claims, which we presume others will share.

There has been increasing awareness of the importance of such issues ever since Watts, Ide and Ehlich (1992: 3) drew a distinction between first-order politeness and second-order politeness:

We take first-order politeness to correspond to the various ways in which polite behaviour is perceived and talked about by members of sociocultural groups. It encompasses, in other words, commonsense notions of politeness. Second-order politeness, on the other hand, is a theoretical construct, a term within a theory of social behaviour and language usage.

On this view, we can study *impoliteness* and *aggression* as first-order lay terms, or alternatively we can propose second-order, technical definitions of “impoliteness” and “aggression” for the purposes of scientific analysis and theorisation. As long as scholars maintain a clear distinction between the two there should be little problem.

However, this begs the question of what are the grounds for the technical definitions we propose? In other words, what is the relationship between our scientific programs of study and the social behaviours that are the objects of that study? This question has exercised (im)politeness scholars for more than two decades. Much of this debate has arguably arisen due to the first-second order distinction taking on a range of different understandings. Culpeper and Haugh (forthcoming) propose three broad groupings into which these different understandings fall:

(1) commonsense, ordinary, non-academic versus academic, technical or scientific ways of talking and thinking about (im)politeness; (2) the understandings of participants themselves versus the understandings of observers of interactions; and (3) emic understandings of (cultural) members versus etic understandings of non-members.

While these contrasting perspectives are clearly related, they are not synonymous (Haugh 2012). Drawing a distinction between commonsense and academic understandings is therefore not as straightforward as might first appear for a number of reasons.

First, we cannot simply equate commonsense understandings with how words like *impoliteness* and *aggression* are used in everyday discourse, and academic understandings with how they are used (and sometimes defined) in scientific discourse. Unless terms are explicitly defined by academics they are arguably being used in a first-order, lay sense. Just because scientists are using a particular term does not mean that term is being used in a scientific manner. Second, we are not only interested in the meanings of words (how people talk about the phenomenon in question), but in underlying concepts (i.e. how people think about that phenomenon). A key complication here is that we cannot equate concepts with the attested meanings of words (Haugh 2016). Third, we are not only interested in how people talk and think about the phenomenon in question, but perhaps more centrally in what people do. The relationship between what people say they do or expect others to do, and what they actually do is itself complex, as studies in pragmatics have repeatedly attested (Verschuere 2000). Studying how “members of sociocultural groups” talk about *impoliteness* or *aggression* clearly offers us a window into how they conceptualise those phenomena, but does not exhaust what we as (im)politeness scholars need to be examining. The technical definitions we propose for the concepts we use are therefore critical as they carve out the scope of what counts as the object(s) of analysis, and, importantly, what does not. They go to the very heart of our warrant for undertaking research in the first place.

Janicki (2017), however, argues that impoliteness and language aggression scholars are needlessly worried about definitions, and that “working definitions” are all that is required. His example is a working definition for the notion of an insult: “to say something that offends” (2017: 162). Note that this is an example of a working definition rather than a working definition constructed as an actual part of a research project on insults, and thus it is

hardly fair to bring the full force of critique to bear. However, we think it is usefully constructive to examine this working definition as if it were a stepping-stone on the path of an insults research project. The key issue with this working definition would be that it begs a further question: what does “offend” mean? Janicki (2017: 162) acknowledges that in his working definition ‘offend’ “still refers to something blurry and very abstract”. However, the presumption that the meaning of “offend” is readily transparent is something that is all too often made in the literature. OFFENCE is a central concept in impoliteness, aggression and conflict research.¹ Yet, despite being widely mentioned in the literature, little by way of systematic study of it has taken place. The notion of offence is often part of definitions of impoliteness, aggression or conflict, but is not actually defined itself. For example, Culpeper (2011: 23) in defining the notion of impoliteness adds that “such [impolite] behaviours always have or are presumed to have emotional consequences for at least one participant, that is, they cause or are presumed to cause *offence*” (emphasis added), but does not clarify what constitutes offence itself. Indeed, it is difficult to find any study to date in the field that really does so in any great depth. While avoiding infinite regress in definitions, that is, defining the words that are used to define words, is clearly one motivation for not trying to define every word we use, OFFENCE is something that deserves, we believe, more concerted attention by researchers in the field.

In this paper, we take Janicki’s (2017) thought-provoking critique as a launch pad for advocating a more considered approach to conceptual and definitional issues in studies of impoliteness, aggression and conflict. We begin by first briefly summarising the key arguments made by Janicki (2017), before going on to extend the discussion with reflection on what scholars of impoliteness, aggression and conflict are doing. We next suggest that a corpus-based approach to metalanguage offers a more empirically satisfactory approach to addressing definitional matters in the field. We illustrate our arguments through a case study of OFFENCE, focusing, in particular, on the use of the term *offensive* in British English in the two-billion word Oxford English Corpus, also noting whether there are particular differences in Australian English. We conclude that the similarities and differences observed therein offer more empirically sound grounds for developing a definition of OFFENCE, as well as raising intriguing new research questions about the extent to which understandings of seemingly routine concepts can vary across communities of speakers.

2. Definitions in research on aggression and conflict

2.1 Janicki’s position

Prof. Karol Janicki is a distinguished scholar, with a track-record of research on definitions (see especially Janicki 1999). Janicki (2017) addresses head on the issue of definitions in research on the language of aggression and conflict. His writing was prompted by the IPRA conference panel (Antwerp, July 2015) in which the editors of the *Journal of Language Aggression and Conflict* (JALC) organized a panel called ‘Researching and Understanding the Language of Aggression and Conflict’. Our aim here is not to engage Janicki’s general approach, one which broadly follows the general semantics of Alfred Korzybski, but to examine some of his specific statements about definitions in the context of pragmatics, and especially in the context of research related to language and aggression, including impoliteness.

Janicki’s (2017: 157) initial observation is that language, aggression and conflict research has a problem and it is one that vexes participants:

¹ We use small capitals here to signal reference to an abstract concept rather than the terms that might realise it.

I have been drawn to the lengthy discussions often expressing dissatisfaction and worries concerning the definitions of the central concepts such as conflict, aggression, hate speech and insult, concerning the somewhat less central ones perhaps, such as impoliteness, politeness, trolling, heckling, irony, etc.

In particular, he notes that it is the lack of agreement about definitions of terms/concepts that causes “deep concern” (2017: 158). These observations, Janicki points out (2017: 158), are not confined to notions related to language, aggression and conflict, but even include the term *pragmatics* itself. In the rest of his 2017 paper, Janicki argues that pragmatics researchers, and more specifically language, aggression and conflict researchers, articulate these worries and concerns about definitions because of their own view of what definition is. Thus, in a sense, it is a problem, he argues, of their own making.

Janicki’s (2017: 159) first point is to challenge the impression that linguists give that they are “particularly unfortunate, to have to face terms which are so difficult to define”. He argues that definitional problems are “independent of the discipline” (2017: 159), and instead “spring mainly from the author’s philosophical position”. In a nutshell, he argues that if we are thinking that we do not have a good definition yet, we are thinking that one ultimate and precise definition is possible; in other words, we are taking an essentialist viewpoint, something which many scholars, from Popper (1945, 1979) to Wittgenstein (1958) and indeed Janicki (2006, 2010, 2015), have argued against. Indeed, Janicki (2017: 160) argues that the attempt to come up with some ultimate definition of the term “will always lead to an infinite regress of definitions (Popper 1945, 1979)”. What this means is that we will need to define the words used in our definitions of a particular term, and the words used in those definitions, and so on, with the result that eventually we may end up using the term we are trying to define originally, and thus we will have made a “full definitional circle” (2017: 160). Such attempts, then, are doomed to failure.

Janicki (2017: 161-4) suggests how we can get to a “rewarding solution” (2017: 161) for the above “problems”. The first step is to adopt a “working definition” of whatever concept we are dealing with in our research or paper, and importantly that definition “must be *useful* for our purposes” (2017: 161), helping solve the problem we are tackling in our research. A working definition, of course, does not assume that it is the ultimate definition, and thus is non-essentialist. The second step (2017: 162) is to “lower the level of abstraction the words in our definition involve”, thereby making it easier “for us to know what we are talking about”. Janicki illustrates with various definitions of *to insult*, suggesting that the definition “to *say* something that offends” has the merit of being restricted to ‘saying’, rather than ‘saying’ and other forms of behaviour, thus becoming less abstract. Janicki (2017: 163) cautions that this definition of *to insult* is to illustrate a definition that is less abstract, and adds that it might not “be useful at all” – “it is up to the individual researcher to decide what definition best suits his or her individual project”. Janicki (2017: 163) comments “in passing” that approaching meanings and definitions with Cognitive Linguistics in mind can be helpful, because it accommodates fuzziness. Thus, with prototype theory in mind, we could talk about typical cases of ‘aggression’, where we would tend to agree that it is a case of ‘aggression’, or less typical cases, where we would tend to disagree.

Finally, in the penultimate paragraph of his paper, Janicki (2017: 164) acknowledges particular “shortcoming” of his “working definition” approach, namely, that it “leads to a long array of definitions which are not compatible [...] And so comparing research results is thus made more difficult”. However, he adds that despite the fact that “the incompatibility of definitions seems true”, we do in fact work with different definitions anyway, even if those differences are small.

2.2 Drawing out issues for aggression and impoliteness scholars

Janicki's (2017) discussion of definitions in the context of language, aggression and conflict helpfully puts focus on pertinent issues. Like Janicki, we think that drawing inspiration from Cognitive Linguistics, and specifically prototype theory, is a useful way of approaching concepts, not least because it allows for gradience, something which is central to pragmatics (we will refer to prototype theory in Section 4.2.1, as part of our theoretical approach to word-meaning). Nevertheless, for a number of reasons the proposals regarding definitions are not entirely satisfactory.

Whilst it absolutely is the case that pragmatics researchers often agonise over definitions and incorrectly assume that defining concepts is much worse for them relative to concepts in other disciplines, we are less comfortable with the idea that this automatically means that they are decidedly working towards essentialist definitions. Janicki (2017: 159) illustrates his argument with a comment from the language and conflict panel at IPRA: "we don't know *yet* what aggression is", which presupposes, of course, that there is a possibility that one day we will. However, this quotation does not help us understand what to know something actually means. Moreover, when people communicate, they draw on common ground to affect economies of expression. Perhaps "know" here means to "know something well enough for the purposes of our research". And if this is the case, this is now close to Janicki's notion of a "working definition" (i.e., "the definition we adopt must be useful for our purposes", p.161).

Aside from usefulness, the other requirement of the working definition approach is that it tries to be "as little abstract as possible" (2017:162). However, the conundrum here is: what happens if what you want to define is abstract, as indeed many pragmatic concepts are? The example Janicki (2017: 162) puts forward for "insult", "to say something that offends", is not convincing, and indeed Janicki (2017: 162) specifically acknowledges the area of difficulty: "It should be added, of course, that what 'saying' means is not entirely clear either and can also be contested" (p. 162). It is true that "saying", assuming we mean "uttering" the insult,² is less abstract than, for instance, meaning, i.e. "to mean something that offends". However, "to say something that offends" does not avoid the abstract issue of meaning, as meaning is part of offends, and "that offends" is a restrictive relative clause attending "saying". The relevant aspect of meaning in "offends" might be described as the perlocutionary meaning (Austin 1962), orienting to the effects of what was said on the target. But what exactly "offends" means is complex, as we demonstrate in Section 4. So, in fact, what "insult" defined as "to say something that offends" actually does is to shift the definitional load to the word "offends"; it does not fully alleviate the problems.

Of course, as we noted above, Janicki (2017: 163) is careful to stress that he is not suggesting that he has coined a definition of to insult that will "be useful" for all, adding that "it is up to the individual researcher to decide what definition best suits his or her individual project". This point, in fact, hints at a weakness in Janicki's approach: it is pitched with the individual researcher in mind. In fact, over recent decades, more and more research is being conducted in teams. Indeed, this is to say nothing of communities of researchers, such as the very one that Janicki encountered at the IPRA conference. If we operate in isolation, we risk failing to capitalise on the work of each other. An illustration is provided by Spencer-Oatey's 1996 paper showing how most scholars have been defining the sociological notions of power and social distance in completely different ways, with the consequence that what we thought we knew about power and social distance turned out to be rather misguided. Only a shared

² *Saying* in English has at least two relevant senses here, to 'utter' something and to 'mean' something (Bach 2001).

and sufficiently precise definition allows for triangulation and replication of studies; without it, we can easily be shooting in different directions. Needless to say, to state that for collaborative research a definition needs to be “shared and sufficiently precise” does not mean it will be exactly the same for all. The code model of communication whereby words and expressions convey meanings into our heads in a straightforward fashion has long been discredited (not least by Sperber and Wilson’s [1986] relevance theory). Even if the wording of a definition is the same for each reader, their interpretation of it will differ, though with a precisely defined definition that variation will be minimized and thus of little consequence for its usefulness.

One thorny issue raised by Janicki is that of definitional infinite regress: terms can be defined by terms that need to be defined by terms. Moreover, further down the line we might end up using the term we originally started with. There are, however, ways of arresting that regress. One solution is to stop the infinite regress at the point where the definition is deemed adequate for the purposes in hand. This would be consistent with Janicki’s working definition approach, though, as we suggested above, it is also consistent with what many people are probably already doing in definitional terms even without thinking of it in terms of a ‘working definition’. Another solution is Wierzbicka’s (e.g. 1972) Natural Semantic Metalanguage (NSM), which claims to avoid this kind of problem by creating an irreducible core of elements – basic ‘semantic primes’. While the NSM model casts light on many issues, the pursuit of semantic primes, which are claimed to be a natural set, ultimately needs to be realised in ordinary language; yet any language will have its own contextual, cultural interpersonal colourings, with the result that the primes can hardly be said to be straightforward invariant primes (see Blumczyński 2013, for a brief critique).

In this paper, we will propose another solution, one that considers metapragmatics and metalanguage as a way of circumventing the problem of infinite regress, at last to an extent, and indeed of helping to contribute to a solution for some of the other problems mentioned above.

3. Metapragmatics and metalanguage

The etymological meaning of Greek-derived *meta-*, meaning ‘with’, ‘between’, ‘after’, ‘beyond’, is not at first blush the most helpful guide to our concerns here. However, the notion of ‘beyond’, as seen in for example *metaphysics*, does allude to the fact that we are dealing with a level of meaning beyond the one to which the prefix is attached. More helpful is the meaning it takes on in epistemology, and indeed in many everyday uses (cf. the entry in the Urban Dictionary), where it means ‘about’ the thing to which it is attached (e.g. *metapragmatics* is about pragmatics, and *metalanguage* is about language). With regard to academic fields of study, this meaning works somewhat better for metalanguage than metapragmatics. For instance, it fits Jakobson’s (1960: 356) famous description of the functions of language and specifically his use of “meta-” in his labelling of the “metalinguistic function”: the use of language to “gloss” language (as one would find in a dictionary, for example). The key focus of metapragmatics, however, is different.

Metapragmatics focusses on reflexive awareness, that is, awareness on the part of users and observers of the pragmatic features of specific uses of language, and the potential meanings they have in context (see, for example, Kádár and Haugh 2013: chapter 9; Culpeper and Haugh 2014: chapter 8). As Hübler (2011) points out, sometimes the focus is on the *use* of explicit expressions (e.g. “Is that a *promise*?”, “Don’t be *rude*”); at other times it is on implicit indicators of metapragmatic awareness in their articulation of beliefs and understandings (including, for example, the role of indexicals; cf. Silverstein 2003). Something of a cline of explicitness is apparent in the four-way categorization of reflexive

metapragmatic awareness proposed by Kádár and Haugh (2013: chapter 9). Essentially, “metalinguistic awareness” focuses on the expressions in a language that represent concepts, including pragmatics-related concepts (e.g. *promise*, *rude*); “metacommunicative awareness” focuses on the “interpretations and evaluations of social actions and meanings” (2013: 186), especially as articulated through explicit comments (e.g. “Is that a *promise*?”, “Don’t be *rude*”); “metadiscursive awareness” focuses on social discourses, such as “talk about *politeness* or *courtesy* in the popular media” (Kádár and Haugh, 2013: 187); and finally “metacognitive awareness” focuses on mental states, “presentations of cognitively grounded states, such as attitudes, expectations and so on, through discourse or pragmatic markers” (2013: 187).

Our study of terms related to the notion of OFFENCE focusses on the more explicit aspect of metapragmatics, more specifically, the metalinguistics of impoliteness. As in many languages, the English lexicon is replete with words such as *impolite*, *rude*, *discourteous*, *aggressive* and *ill-mannered*. A focus on explicit metalanguage is not an alien enterprise for some (im)politeness scholars. For example, Eelen (2001) and Watts (2003), claim that “pseudo-scientific” classic politeness theories (e.g. Brown and Levinson 1987) ignore the layperson’s usage of politeness terms and what they might mean, and moreover miss the opportunity to put (im)politeness studies on a firmer ontological footing. However, in practice those particular scholars undertook little in the way of actual study of politeness terms. Such studies have gathered momentum, however, since that time (see Kádár and Haugh 2013: 188-194, for an overview). Especially relevant to our concern here are studies of impoliteness-related expressions. For instance, an early study is Culpeper (2009; see also 2011: Chapter 3), which used corpus analyses to show that *impolite* is not completely synonymous with *rude*, but more precisely matches a subset of its meanings (all with the exception of sex-related ones). Waters (2012) focuses on the cultural semantics of *rude*, using corpus techniques to investigate the formula “It’s rude to VP, and the fixed collocation “rude word” through Google searches restricted to Australia. Needless to say, studies are not restricted to English. Of particular note are the works of Taylor (2015, 2016, 2017), which have pioneered cross-linguistic/cultural work. Her studies show, for example, that English *sarcastic* and Italian *sarcastico* do not occupy the same semantic space, the English term being associated with behaviours that are more negatively evaluated. It is of no surprise that metalinguistic studies often deploy corpus methods. Corpus methods are the methods of choice for the lexicographer mapping the semantic space of words. Further, corpus linguistics maps meaning through a consideration of actual use. This emphasis suits an enterprise rooted in pragmatics.

Why might a study of metapragmatics – specially metalinguistics – particularly suits our endeavour of understanding the concept of OFFENCE better? Firstly, there seems to be an intimate connection between metalanguage and impoliteness. Cameron (2004: 313) points out that morality is an important theme in studies of metalanguage: ‘[m]etalinguistic resources seem very often be deployed to connect various aspects of linguistic behaviour to a larger moral order’. This is particularly pertinent to impoliteness, as instances of impoliteness-related metalanguage are often occasioned by perceived breaches of morality, that is, beliefs about how things ought to be clashing with a perception that they are not.

Secondly, and more generally, examining impoliteness-related metalanguage, both the specific expressions that are used and how they are used, gives us a window into the frames of understanding that people themselves are using. The link between metalinguistic representations and ways of thinking, especially ideologies, has been pointed out by Coupland and Jaworski (2004: 36-37) and Verschueren (2000: 450-452), among others. The systematic metalinguistic study of such terms is therefore not merely an exercise in lexical

semantics (Terkourafi 2005: 242; cf. Terkourafi 2011), as their meaning goes to the very heart of social reality across languages and cultures.

Thirdly, and of particular relevance to this paper, investigating the notion of OFFENCE through corpus-based studies of metalanguage provides an empirical solution to the problem of individual researcher bias. The meaning of OFFENCE is characterised through the use of terms that represent it. Capturing the use of terms is exactly what the corpus method is designed to do, as it allows us to examine subtle gradations of meaning, meanings coloured by variation in community, cultural setting, and so forth. One might object that such meanings still need to be described by other words, and so are still subject to the problem of definitional infinite regress. However, as we elaborate in Section 4.2.1, what underlies the corpus-based approaches to word meaning in particular is examining associated words, that is, collocates. Words are defined by the company they keep (cf. Firth 1957: 11). Thus, rather than the researcher selecting words to define a word in isolation, the objective is to identify the words that co-occur with the target word, and examine, in turn, how these are systematically related to those co-occurring terms in dynamically constituted “semantic fields” (Trier 1931; Ullmann 1962). When that objective is met, there is no need for further regression, as it is the *position* of the word in question and its syntagmatic and paradigmatic *relationships* with other words within that field that in effect defines it.³

4. A study of OFFENCE and its meaning

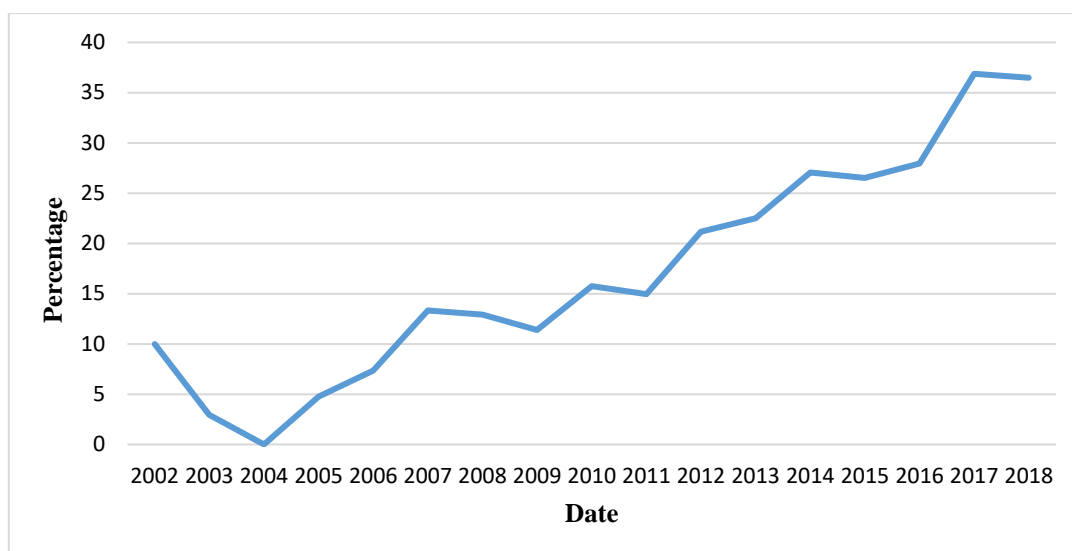
4.1 Some observations on the notion of OFFENCE in research on linguistic impoliteness and taboo language

Our decision to study offence is mainly motivated by its increasing importance in the impoliteness literature.⁴ The notion of OFFENCE is central to research on linguistic impoliteness and taboo language. This was not always explicitly deemed so. Culpeper (2011: 19-20) presents nine quotations from the linguistic impoliteness literature, all defining impoliteness, and a further four definitions of related phenomena from the literature on aggression. None of them mention OFFENCE explicitly. However, that concerns a tendency in research prior to 2011. It was around 2011 that more prominence was begun to be given to the notion of OFFENCE. Evidence of this tendency can be seen in Figure 1, which displays the proportion (expressed as a percentage) of the 2,218 documents published between 2002 and 2018 containing the term *impolite* (or related form) which also contain the term *offence* (or related form), as retrieved by the search engine SCOPUS.

Figure 1. The proportions of documents containing both *impoliteness*-related terms and *offence*-related terms published between 2002 and 2018 out of the total numbers of documents containing just *impoliteness*-related terms for each year

³ A syntagmatic perspective examines collocates of the word in question found in co-text, while a paradigmatic perspective examines recurrent patterns in how the word in question is co-determined by other semantically related words (Hübler and Busse 2012; see Haugh 2018, 2019 for further discussion).

⁴ It is important to bear in mind here that use of the use of the term *offence* (and related terms) in the scientific literature does not mean it is being used in a technical, scientific (second-order) sense. Indeed, given it is rarely defined as go on to point out, we must assume that it is frequently being used by scholars in its ordinary, lay (first-order) sense.



What Figure 1 clearly shows is a notable increase, from around 10% to around 36%, in the explicit mention of OFFENCE in impoliteness-related publications after 2011.⁵

In Culpeper's (2011:23) own definition OFFENCE clearly figures:

Impoliteness is a negative attitude towards specific behaviours occurring in specific contexts. It is sustained by expectations, desires and /or beliefs about social organisation, including, in particular, how one person's or a group's identities are mediated by others in interaction. Situated behaviours are viewed negatively – considered “impolite” – when they conflict with how one expects them to be, how one wants them to be and/or how one thinks they ought to be. Such behaviours always have or are presumed to have emotional consequences for at least one participant, that is, they cause or are presumed to cause *offence*. Various factors can exacerbate how *offensive* an impolite behaviour is taken to be, including for example whether one understands a behaviour to be strongly intentional or not. (our emphasis)

These two references to OFFENCE encompass both a speaker's perspective on communication and a hearer's, that is to say, a speaker causing offence and a hearer taking something as offensive. But what exactly does offence encompass here? It is clear from the quotation that offence is aligned with emotional responses. At various points in Culpeper (2011), the expression “offensive language” is used, and much of Chapter one is devoted to a discussion of “offence types” (i.e. breaches / infringements of sociality rights and quality/identity/relational face) and the first sections of Chapter two examine intentionality and emotion in relation to offence, but nowhere is offence actually defined. Of course, one might say that by not defining offence in this definition, we avoid infinite regress, but it cannot be denied that this is a lacuna.

Haugh (2015), in a paper focused on “impoliteness” and “offence” in initial interactions, points out that taking offence is much more than an emotional reaction triggered by certain actions. He elaborates:

taking offence is analysed as a social action in and of itself distinct from any feelings of offence a participant may or may not experience. This means that those persons registering and sanctioning offence in a particular interaction are not only holding

⁵ Before 2010, raw frequencies are less than 10. For this reason, not much can be inferred about the dip in 2004.

another person (or group of persons) accountable for causing offence, but can themselves be held morally accountable for this taking of offence. (Haugh 2015: 37)

However, whilst the important point that the taking of offence involves “both a moral stance and an affective stance on the part of the producer” (Haugh 2015: 41) is well made, we still lack a sense of what OFFENCE actually is. Haugh and Sinkeviciute (2019) point out in an overview of literature to date that OFFENCE is multifaceted, encompassing a range of different senses. These include OFFENCE as a form of transgression (that is considered *offensive*), the act of transgressing (whereby one *offends*), and the feelings of hurt, indignation and so on engendered by that transgression (namely, where some party is *offended*). They also draw attention to the need to distinguish between causing OFFENCE and taking OFFENCE. The latter can refer to either the feelings that accompany OFFENCE, or a morally accountable action in its own right, or both. The greater focus on the hearer perspective and on “offence-taking” in recent years probably largely accounts for the rise we see in Figure 1.

Perhaps the key link between impoliteness studies and studies of taboo language is precisely the notion of OFFENCE. This is clear in Keith Allan’s introduction to the impressive, cutting edge 2019 collection of papers on taboo words and language. Allan (2019: 2) writes that “to engage in tabooed behaviour is to cause offence to others and so it is dysphemistic”; “dysphemism is typically impolite because it is offensive”. Note here that the phrasing in these quotations suggests a speaker-centred perspective of causing offence. However, the book encompasses much more than this. For example, Jay’s (2019) chapter encompasses the perception of the offensiveness of taboo words, whilst Dewaele’s (2019) chapter encompasses the perception of taboo words in proficient adult foreign language users, including specifically offensiveness. Nevertheless, rather like earlier impoliteness research, there is little focus on the taking of offence. Also, like impoliteness research, but this time in general, there is no attempt to tackle head-on what being “offensive” might mean, though there is plenty of discussion about what it involves.

The rise of explicit mention of offence in linguistic impoliteness research, which we observed above, is not reflected in taboo language research. Figure 2 displays the proportion (expressed as a percentage) of the 11,511 documents published between 2002 and 2018 containing the term *taboo* (or related form) and either the term *linguistic* or *language* (or related forms), and which also contain the term *offence* (or related form), as retrieved by the search engine SCOPUS.

Figure 2. The proportions of documents containing both *taboo*-related terms and *linguistic*- or *language*-related terms and also *offence*-related terms published between 2002 and 2018 out of the total numbers of documents containing just *taboo*-related terms and *linguistic*- or *language*-related terms for each year

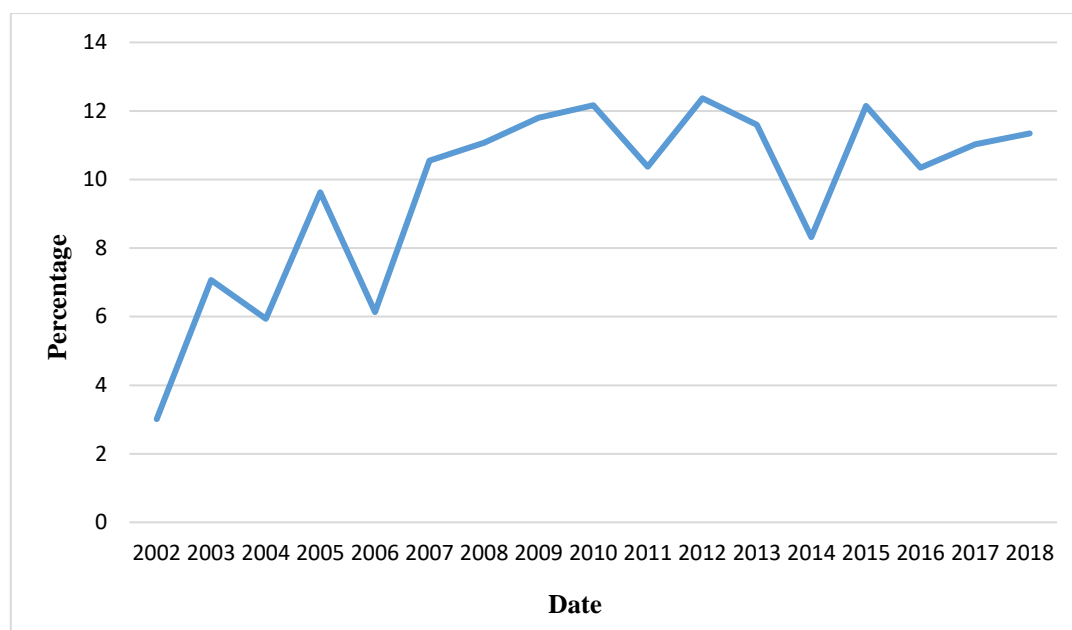


Figure 2 does display a slight upwards trend up to around 2007, but then it seems to plateau. We see a much flatter line than that of Figure 1. In fact, if we exclude the first year, 2002, percentage values only vary between 6 and 12%. There are probably several, complex reasons as to why research on taboo language does not mirror the rise of offence-related terms that we saw with impoliteness research. For one, the study of taboo language infuses many fields of study that are not overly concerned with how hearers experience or take offence, but with how, for example, particular behaviours become or are treated as taboo in particular communities (cf. anthropological, historical, sociological studies). It is true that the scholars mentioned in the previous paragraph – Allan, Jay and Dewaele – do clearly consider the offensive effects of taboo language, but all three are linguists, and thus not entirely representative of the entire domain in which taboo language is discussed, often fleetingly, by scholars of many hues. An additional factor is likely to be the fact that the whole notion of taking offence in language interaction is a driver for linguists working on impoliteness, but does not figure in studies of taboo language, even some studies involving linguists.

4.2 The metalinguistics of OFFENCE: A corpus-based study

4.2.1 Our approach and focal point

Our aim in this section is to use corpus-based methods, the methods of choice for the modern lexicographer, to define the term *offensive*, a term that is used to denote things deemed to involve OFFENCE, and to explore how this might inform scientific analysis and theorisation of it. It is worth drawing attention to underpinning tenets behind the corpus linguistic approach to meaning. First, the corpus linguistic approach is a usage-based approach to meaning. Traditional dictionary makers or lexicographers also look at the use of words. However, their approach is often top-down, looking for evidence of meanings, rather than bottom up, letting patterns of use determine meanings. Second, meanings involve frequency of use. Anybody can invent a meaning for a word, but it would be a one-off nonce word. Word-meanings are conventions accepted by particular speech communities, and established through the repetition of the word in similar contexts. Third, words and their meanings are sensitive to the words with which they regularly co-occur, i.e. their collocates (cf. Firth 1957). Collocates will be a central part of our method. Fourth, and partly as a consequence of the third, word-

meanings are bound up in broader chunks of language than single, isolated words (cf. Sinclair, e.g. 2004). Fifth, and partly as a consequence of the fourth, grammar is not separate from lexis: those chunks of language are made up of words that participate in certain grammatical patterns, and that participation is part of the conventional meaning of words (cf. Sinclair, e.g. 2004). We will examine the particular collocates of *offensive* in specific grammatical patterns. This is a crucial point: identifying those lexico-grammatical patterns is a way of teasing out different meanings. Relevant to this, note the point made by Goldberg (1995: 67), one of the seminal works on construction grammar: “if two constructions are syntactically distinct, they must be semantically or pragmatically distinct”.⁶

Of course, teasing out the different strands of meaning does not address the issue of what holds everything together. We would argue for a cognitive theory, specifically, prototype theory (e.g. Rosch 1975). Corpus linguistic analyses are highly compatible with cognitive theories. Both, for example, involve experience and frequency (for further elaboration on the connections between corpus linguistics and cognitive linguistics, see Schönefeld 1999). Prototype theory constitutes an alternative approach to Aristotelian categories based on necessary and sufficient conditions. Instead, it is based on typical features as acquired through experience. For example, for the category “bird” the ability to fly might be taken as a necessary feature, but that would exclude an ostrich, which nevertheless has some bird-like features (e.g. feathers). Prototype theory offers proper theoretical status to gradience, and takes account of the varied nature and complexity of criteria for category membership (cf. Coleman and Kay 1981). Importantly, it follows that what we are aiming at in this paper at is not a traditional dictionary entry. We are especially not attempting to echo the kind of entry found in the Oxford English Dictionary, where we find etymologically-ordered ‘senses’ imposed by an editor to organise the material at their disposal.⁷ We are interested in using corpus-methods to reveal aspects or features of meaning that are regularly experienced – they are part of the prototype. We refer to these aspects or features of meaning as ‘semes’ (from Greek: *sēma*, ‘sign’, a minimal distinctive unit of meaning). Crucially, when reading our definition it should be remembered that it represents the *probable* features of meaning of offensive, not the *possible*.

Our aim in all this is not to produce a definition of *offensive* that is fixed for all time and contexts, but to show how a particular method can be used to provide some empirical anchorage for the notion of OFFENCE in a specific context, namely, various written discourse genres of English produced in Britain in the early 2000s. Needless to say, we are not denying that more fine-grained detail is impossible. Fortunately, statistics allow us a means of controlling granularity: less salient features will not, for example, be captured by the association measures we set in place. There is no gold standard for the right amount of detail; it is often a matter of fitting the purposes in hand. Practical constraints, such as the amount of space for the definition and the stamina of the readers, are usually important.

One might ask: why study the adjectival term *offensive* in particular? There are indeed other morphologically related words, such as *offend* or *offence*, that we might have studied instead (or in addition, if we had space). It may seem the obvious choice to pick the most

⁶ As Maarten Lemmens kindly pointed out to us, this dictum permeates much scholarship. Weirzbicka (1988: 3), for example, states that “every grammatical construction encodes a certain meaning”. In fact, her entire book is an articulation of the idea that there is no division between grammar and semantics, or pragmatics, for that matter. Interestingly, this reverse of this dictum – a difference of meaning leads to a difference of form – is not true (see, for example, Croft 2001: 76).

⁷ A large proportion of the Oxford English Dictionary has not been revised since its construction in the late nineteenth-century / early twentieth century. Regarding *offensive*, the entry does have one citation from the twenty-first century, suggesting that the editors, currently working on the third edition, may have reached this entry. However, one citation for the twenty-first century does not inspire confidence in their treatment of meanings in recent English.

frequent item, namely, the noun *offence* (121,413 instances in the two-billion word Oxford English Corpus (OEC)). However, overwhelmingly, *offence* is used in legal contexts; it is thus less instructive about what OFFENCE generally means. The adjective *offensive*, with 46,188 hits in the OEC, clearly occurs more frequently than the verb *offend*, with 34,671 hits. Frequency is important not just to enable more robust results, but because frequency of use reflects frequency of experience, and thus *offensive* plays a more central role in shaping the concept OFFENCE (and in turn being shaped by it).

4.2.2 Method

For our method to work, we need a corpus of considerable size. The version of the *Oxford English Corpus* (OEC) we used consists of more than two billion (2,073,319,589) words of English, largely culled from the web between 2000 and 2006. What is key to our study is that, amongst other categories, the data is coded for “language type”, that is, the variety of English (Cook and Brinton 2017). Our study is primarily based on the British English component (502,259,374 words), and secondarily the Australian English component (89,328,901). Our choice here is influenced by the fact that the authors of this study are members of these cultures, and they are amongst the largest data components in the corpus (the fact that the British component is the primary focus of our study is down to its size, something which affects the richness of the results).

Our analyses of the OEC were performed by *Sketch Engine* (SE).⁸ SE was largely designed with lexicographers in mind (see Kilgarriff et al. 2004, for an important paper describing it). This is entirely appropriate to our enterprise here, as word-definitions are the central part of what lexicographers do. There are numerous corpus tools that can be used to examine the co-occurring words of a target word (often within a certain span of that item, such as five words prior and five words afterwards), and identify, through statistical measures, those that have a strong association (i.e. its collocates). What SE does over and above this is to search for collocates within the grammatical relations that the target word participates, an analysis which is referred to as a “word sketch”. For example, in OEC British English, with the target word *bastard*, the top collocates modifying *bastard* as a noun are *lazy*, *lying*, *miserable*, *fat* and *evil*, whereas the top collocates of *bastard* modifying a noun are *offspring*, *lovechild*, *feudalism*, *son* and *cousin*. These grammatical relations give us some insight into the contextual relations that the term participates in – we learn about the who, what and how of *bastard*. SE examines a target word for 27 grammatical relations, and uses the statistic logDice (Rychlý 2008) to assess the strength of the collocates and rank-order them.

In this case study, we first conduct a Word Sketch of *offensive* in OEC British English, and then compare that with *offensive* in OEC Australian English using Word Sketch Difference, which considers whether a word shares behaviours or not. Of course, when analysing language at scale in this way, and thus necessarily deploying automated or semi-automated processes, the usual caveats apply. For example, part of speech tagging errors may occasionally drop an invalid result. This is why it is always important to scrutinise the detail of the output. As far as possible our analyses are data-driven. If we think that *offensive* means something but there is no evidence for it, it will not go into our definition. However, interpreting groups of collocates is somewhat like joining the dots to see the larger strand of meaning, and that act of interpretation will inevitably be coloured by our prior knowledge.

4.2.3 Offensiveness in British English

⁸ <https://www.sketchengine.eu/>.

Table 1 displays the results of a Word Sketch analysis of the term *offensive*. The grammatical patterns were rank-ordered according to the number of collocates they contain, and the collocates in each grammatical pattern were rank-ordered according to their strength of association with *offensive*. We have not included any grammatical pattern that included fewer than 10 collocates. We also excluded: (1) the pattern “offensive + preposition”, simply because the resultant list of collocating prepositions offers little of interest; and (2) the pattern identifying collocating verbs when they are followed by noun phrases of which *offensive* is the adjectival complement, because these results are substantially skewed by the fixed expression “launch/mount/continue/etc. a charm offensive”. For each grammatical pattern we include two examples, along with the overall frequency of collocates in that grammatical relation and its association score (logDice). The examples illustrate the top two most frequent collocates in the pattern. In order to reduce potential bias in our selection of examples, we deployed the GDEX (Good Dictionary EXamples) function in Sketch Engine, a tool that automatically evaluates sentences for length, advanced vocabulary, sufficient context, pronouns pointing outside of the sentence and other criteria, and thereby identifies sentences which are easy to understand and illustrative enough to serve as examples (see <https://www.sketchengine.eu/guide/gdex/>). In the second column of the table, we list the collocates of *offensive* in that grammatical relation, up to a total of 10, along the overall frequency and association score of each.

Table 1. The results of a Word Sketch analysis of the term *offensive* in the OEC British English (performed by Sketch Engine)

<i>Grammatical pattern</i> (frequency/association score) (patterns are numbered for ease of reference)	<i>Top-ten collocates of offensive in that grammatical relation</i> (frequency/association score)
[1] <i>Nouns it pre-modifies</i> (32,848/71.12) “He can also be controlled by bigger <u>offensive linemen</u> ” “He was the <u>offensive coordinator</u> at Oregon from 1998 to 2001”	lineman (1,533/10.37), coordinator (2,184/10.12), tackle (911/9.23), weapon (1,298/8.04), line (4,014/7.98), operation (977/6.93), remark (207/6.88), capability (310/6.62), coach (379/6.53), player (1030/6.41)
[2] <i>Conjuncts with which it shares and/or constructions</i> (10,287/22.27) “Such structures will still affect military <u>offensive and defensive</u> planning” “Some thought applause after religious performances was <u>offensive and inappropriate</u> ”	defensive (1,067/10.67), inappropriate (123/8.08), insulting (99/8.05), obscene (60/7.39), abusive (75/7.37), tasteless (55/7.25), rude (64/7.2), outrageous (60/7.2), insensitive (50/7.1), strategic (128/7.07)
[3] <i>Adverbs by which it is pre-modified</i> (4,883/10.57) “Black people know <u>racially offensive</u> stuff when they hear it.” “it is <u>deeply offensive</u> to both her and me to keep reading about the terrible financial strain his ex-wives were upon Philip.”	racially (127/9.16), deeply (444/8.87), grossly (135/8.87), morally (106/7.63), personally (57/7.6), downright (61/7.34), potentially (165/6.94), mildly (36/6.84), profoundly (35/6.55), religiously (16/6.4)
[4] <i>Subjects of which it is an adjectival complement</i> (2,108/4.56)	comment (37/8), remark (19/7.82), ad (15/7.41), word (25/7.16), term (16/7.11), joke (11/7.04), speech (11/6.67), behaviour (16/6.53), language (13/6.43), movie (14/6.35)

<p>“He later apologized if his <u>comments were offensive</u>.”</p> <p>“Whether true or false, Dershowitz's intemperate <u>remark was offensive</u>.”</p>	
<p>[5] <i>Nouns in “to” prepositional phrases which post-modify it (1,852/4.01)</i></p> <p>“the term is <u>offensive to Muslims</u>, no doubt”</p> <p>“I think he did some things that are highly <u>offensive to the sensibilities</u> of most Americans in general and women in particular.”</p>	<p>Muslims (47/9.15), sensibility (31/8.57), Christian (21/7.81), Catholic (16/7.61), Islam (32/7.53), dignity (14/7.31), non-Christian (8/7.07), scripture (8/6.91), Jew (10/6.9), witch (7/6.82)</p>
<p>[6] <i>Verbs where it is an adjectival complement (750/1.62)</i></p> <p>“Remember that it is what the employee <u>finds offensive</u> that counts.”</p> <p>“Should firms like ClearPlay be allowed to remove content they <u>deem offensive</u>?”</p>	<p>find (323/9.09), deem (34/8.17), consider (39/7.42), smell (9/5.92), sound (35/3.91), seem (82/3.55), prove (16/3.04), become (85/2.71), appear (9/2.58), get (36/1.94)</p>

Table 1 cannot be read as a set of traditional dictionary ‘senses’; indeed, they do not have the look of the descriptive commentary that editors impose on any data they use. However, in tune with our approach outlined in 4.2.1, each grammatical pattern with its associated collocates points to a different seme. Below, we offer a brief interpretive remark on the collocates of each grammatical pattern, starting at the top and working down the grammatical relations, which are numbered for ease of reference. Although we are largely discussing these separately, it should be acknowledged that some of these semes clearly intersect with each other. Together all the semes form a network. These interpretations do not simply arise from the collocates as displayed in Table 1. Each collocate must be inspected for its use. We did this by generating and examining a concordance of each one.

- (1) Collocates, such as *lineman*, *coordinator*, *tackle*, *coach* and *player*, provide strong evidence that, amongst other things, offensiveness is associated with competitive sports and their tactics. We should also mention here that there is a clear, though limited, intersection between this seme and the following one. *Defensive* is the most frequent collocate in pattern (2) involving conjuncts, where it contrasts with *offensive* in this domain specifically. In fact, its high frequency in (2) – 1,067 instances – is much more like the collocates in (1) than (2). Upon inspection, *defensive* shares the same competitive sports contexts as the collocates in (1), and also, frequently, video gaming and military contexts (see the example in Table 1).
- (2) Strikingly, many of the collocates here – *inappropriate*, *insulting*, *abusive*, *rude* and *hurtful* – are mentioned in the impoliteness literature. All these involve negative subjective judgements of behaviours. *Hurtful* identifies a specific emotion.
- (3) Of note is the fact that the strongest collocate is *racially*, and we also see *religiously*. These associations reflect the sensitivities of the culture, especially specific social groups, to particular kinds of offence. But note that *offensive* is also associated with offence relating to individuals – *personally*. Additionally, in the collocates here a scale of offence is evident, from *deeply/grossly/downright/etc.* through to *mildly*, though, significantly, note that the bulk of the discourse is at the high end. This is clearly important information in understanding the meaning of *offensive*, but it is not, incidentally the kind of

information that constitutes a traditional dictionary ‘sense’. Finally, we briefly note *morally* in fourth position: offensiveness is associated with morality.

- (4) Clearly, linguistic behaviours, including specific words or terms, dominate what attracts the attribution of offensive. These are the targets of the negative subjective judgements.
- (5) What is generally being captured here are the people who suffer the offensive entity, typically a linguistic behaviour, as outlined above. The striking pattern in the collocates is religion – *Muslims, Christian, Islam, non-Christian, scripture* and *Jew* (a collocate that is typically used here in reference to religious aspects rather than the people).
- (6) The dominance of private, cognitive verbs emphasises the point that being offensive is a subjective judgement.

These semes form a network that we can weave together into a definition, thus:

Being offensive is a subjective judgement that a behaviour (typically a linguistic behaviour) is insulting, abusive and/or hurtful to, especially, religious groups, ethnic groups or specific individuals, and considered inappropriate, rude and/or morally wrong. The quality of offensiveness varies in intensity, though is usually at the more intense end of the scale. (A separate domain of use relates to competitive sports, especially football, where it refers particularly to the tactics deployed in opposition to an opposing team, and contrasts with defensive).⁹

This definition, though of course it could be reworded in many different ways, is arguably not as subject to the problems raised by Janicki (2017). First, it helps carve out analytical territory for impoliteness and aggression researchers in an empirically motivated way; that is, based on an understanding of the term in question that is aggregated across a community of users rather than being based on the intuitions of the individual researcher (Culpeper 2009; Haugh 2016). Second, it helps us better understand the metalinguistic resources that are systematically available to users of a community when claiming or taking up OFFENCE. Any definition of a word that is based on aggregated understandings of it is, of course, necessarily abstracted away from the particular, often more granular understandings of participants in situated contexts (Haugh 2019). However, crucially, when using words in particular situated contexts we are nevertheless constrained by what other members of that speech community will take as legitimately meant by it, and so understandings of words that are presumed to be shared across communities are an important resource for users.

Our definition here could be enhanced through systematic examination of the collocates of key related terms, such as *insulting, abusive* and *hurtful*, and the relationships between them and *offensive*. This kind of approach can help address, in part at least, the problem of definitional regress by positioning the focal term within a broader semantic field of related terms. The extent to which researchers might systematically examine those related terms depends, in turn, on the particular focus of their study. However, our aim here is not to offer a definitive proposal about what is meant by *offensive* (in British English or any other variety of English for that matter), but to show how corpus-based metapragmatic approach offers a more empirically sound pathway to understanding our objects of study. To illustrate the importance of this approach, the point we would like to pursue in the remainder of this paper is that understandings of OFFENCE can vary across different speech communities and

⁹ Prompted by a reviewer’s comment, we compared our definition with the definition given in the latest edition of what is probably the most famous corpus-based dictionary, the Collins Cobuild English Language Dictionary (available online at <https://www.collinsdictionary.com/dictionary/english>). Two “senses” are supplied for *offensive* as an adjective: (1) ‘Something that is offensive upsets or embarrasses people because it is rude or insulting’, and (2) ‘In sports such as American football or basketball, the offensive team is the team that has possession of the ball and is trying to score’. This is clearly compatible with what we have produced, except that ours is considerably more nuanced.

cultural contexts, and this has consequences for how we go about studying impoliteness and aggression.

4.2.4 Are there differences between British and Australian English conceptions of OFFENCE, and if so, what are they?

In this section, we compare the results of our OEC British English word sketch with results of an OEC Australian English word sketch using the tool Word Sketch Difference. This works by generating two Word Sketches, and then comparing each of their collocates in each grammatical relation. Figures 1 - 3 display visualisations of the Word Sketch Difference results for the three grammatical relations where differences between the British English and Australian English corpora were clearest. The size of the circles and collocate text is proportional to the frequency of the collocate; the colours in the circles containing collocates indicate whether the collocate is more frequent in the British or Australian component; the positioning of the circles horizontally indicates the strength of the collocation, i.e. whether it forms a stronger collocate in the British or Australian components or whether it is somewhere in between. It should be remembered that the British corpus component is bigger than the Australian, and therefore nothing can be concluded from the fact that the collocate circles and text, representing their frequencies, are often larger for the British component, or that there are generally more collocates arising from the British component. We will not comment further on the separate domain relating to competitive sports.

Figure 3. Conjuncts with which *offensive* shares “and/or” constructions: British and Australian English compared

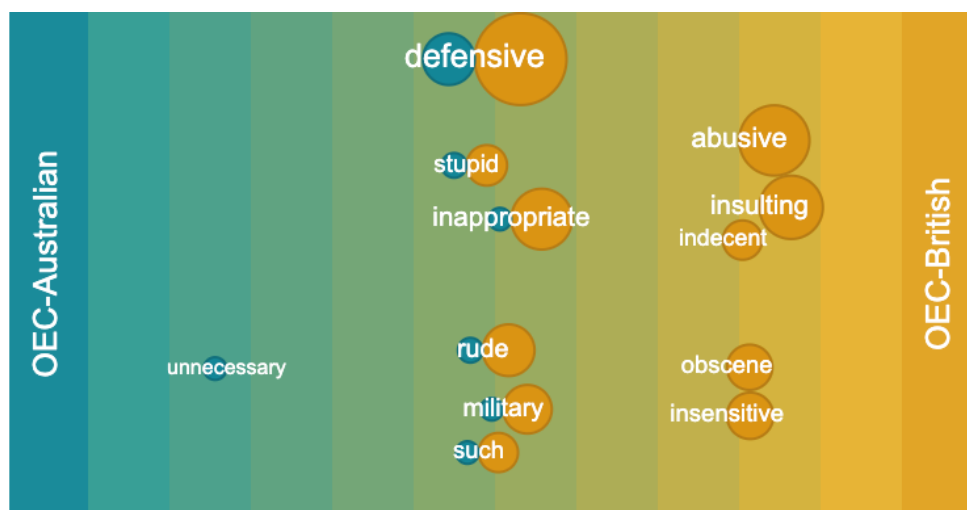


Figure 4. Adverbs by which *offensive* is pre-modified: British and Australian English compared

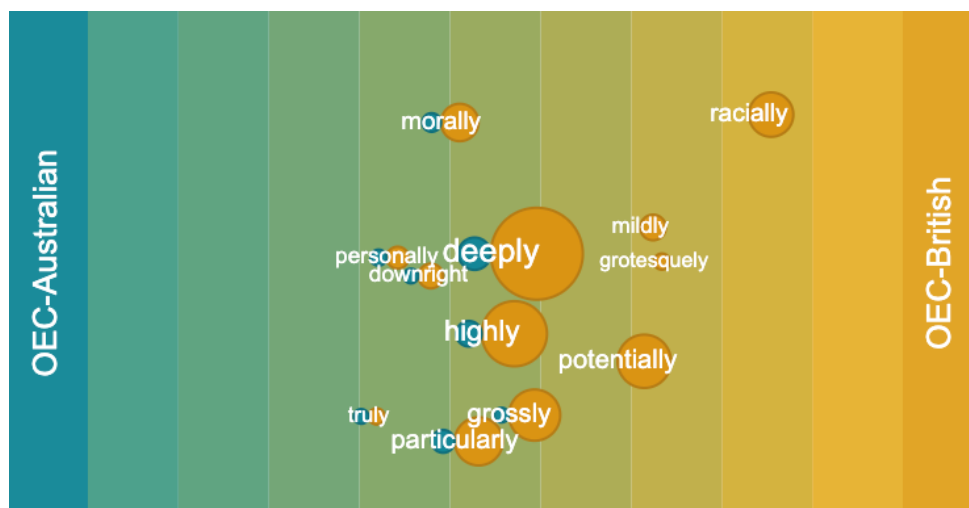
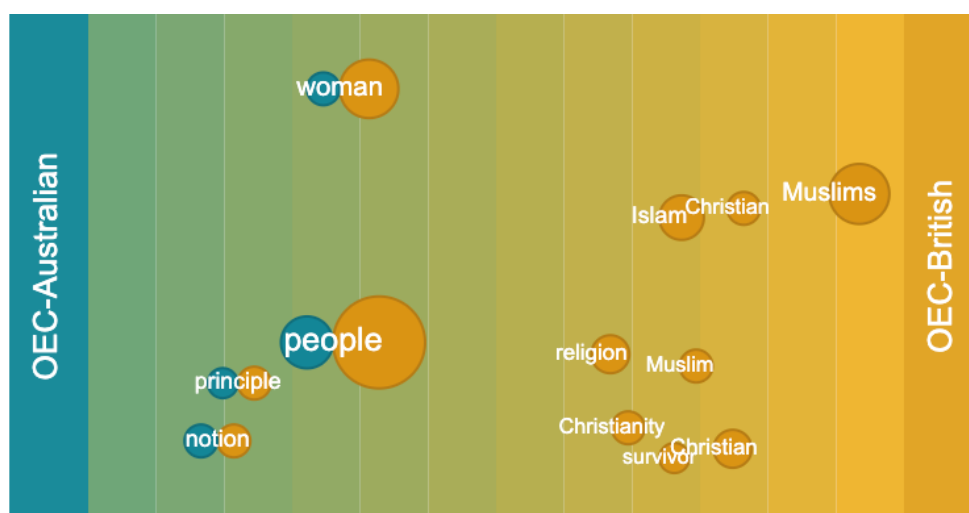


Figure 5. Nouns in “to” prepositional phrases which post-modify *offensive*: British and Australian English compared



As can be seen in Figure 3, the idea that offensive behaviours relate to behaviours that are inappropriate and/or rude is common to both Englishes. What marks British English out is the collocates *abusive* and *insulting*, which were also mentioned in the previous section, are the strongest collocates in British English and amongst the most frequent. This raises the possibility that negative emotional consequences of being *offensive* are closer to the centre of the British conception compared with the Australian (this is also consistent with the presence of the collocate *insensitive*). The collocates in Figure 4 are largely shared. The exception here, however, is the collocate *racially*. Racially motivated offensiveness is a feature of British English, but not Australian English. Finally, Figure 5 captures some distinct differences. The key concern with religiously motivated offensiveness in British English is not reflected in Australian English. Of course, these observations would need to be replicated with a larger Australian corpus, but certainly they suggest some revisions to our definition of *offensive* in British English are needed for a definition of *offensive* in Australian English. A definition might be as follows:

Being offensive is a subjective judgement that a behaviour (typically a linguistic behaviour) is considered inappropriate, rude and/or morally wrong. The quality of offensiveness varies in intensity, though is usually at the more intense end of the

scale. (A separate domain of use relates to competitive sports, especially football, where it refers particularly to the tactics deployed in opposition to an opposing team, and contrasts with defensive).

We would caution that this definition of *offensive* in Australian English, and the earlier one in British English, is informed by analysis of primarily written discourse collected at a particular time (the early 2000s) from the Internet. The meaning of different words can change over time, so it remains an open question whether such differences persist. It also remains open to further study whether such understandings might be different across other more interactive modes of communication, such as spoken or digitally-mediated conversation. For the moment, we suspect some of the key differences we have sketched here between understandings of *offensive* in Australian and British English may persist, but this remains for the moment conjecture.

5. Concluding remarks

In naming and defining things we are thereby doing things, including making (tacit) claims about “social realities”. The definition of words is thus a concern for both scientific communities and communities of speakers. However, the nature of these concerns are different. For the former, definitions are vital for the triangulation and replication of studies and broader collaborative research programmes about key notions such as OFFENCE. This is particularly important in the case of impoliteness, aggression and conflict research which is, by its very nature, necessarily interdisciplinary. For the latter, notions such as OFFENCE are deployed in the course of the moral positioning that (endlessly) occurs in interpersonal and intergroup settings. As we have seen, the understanding of such notions can vary across speech communities. The understanding of analogous notions across different languages – bearing in mind the possibility that a concept encapsulated in one word in one language may be distributed across more than one word in other languages – can clearly vary even further. If we are going to get a proper handle on our object of study across languages and cultures in research on impoliteness, aggression and conflict, we cannot ignore this variability.

Janicki (2017) argues that we need not argue about definitions in the field of impoliteness and aggression research as they are, in the end, only a function of a particular set of research questions. We have suggested that it is rather more complicated than that. The language we use in our analysis defines – that is, it affords and constrains – our object(s) of analysis. Recognising this fact is one thing. Dealing with it in an empirically sound manner is another. We have proposed that one productive way forward is to ground our working definitions (that is, definitions we use as researchers to identify our object of study relative to our research questions) in corpus-based metalinguistic analyses of concepts that are central to our research. In this paper, we have suggested that OFFENCE is one of those concepts. We have also argued that paying due attention to the technical definitions of such terms is not just a sign of navel gazing by researchers. Such things should matter to us as researchers, because it matters to users of language themselves.

References

- Allan, Keith. 2019. “Taboo words and language: an overview.” In *The Oxford Handbook of Taboo Words and Language*, edited by Keith Allan, 1-27. Oxford: Oxford University Press.
- Bach, Kent. 2001a. “You don’t say?” *Synthese* 128 (1/2):15-44.
- Berger, Peter, and Thomas Luckmann. 1966. *The Social Construction of Reality*. NY: Doubleday.

- Blumczyński, Piotr. 2013. "Turning the tide: A critique of Natural Semantic Metalanguage from a translation studies perspective." *Translation Studies* 6 (3):261-276.
- Brown, Penelope, and Stephen Levinson. 1987. *Politeness. Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Cameron, Deborah. 2004. Out of the bottle: the social life of metalanguage. In *Metalanguage*, edited by Adam Jaworski, Nikolas Coupland and Dariusz Galasinski, 311-322. Berlin: Mouton de Gruyter.
- Coleman, Linda and Paul Kay. 1981. Prototype semantics: The English word "Lie". *Language* 57(1): 26-44.
- Cook, Paul, and Laurel J. Brinton. 2017. "Building and evaluating web corpora representing national varieties of English." *Language Resources and Evaluation* 51:643-662.
- Coupland, Nikolas, and Adam Jaworski. 2004. "Sociolinguistic perspectives on metalanguage: reflexivity, evaluation and ideology." In *Metalanguage. Social and Ideological Perspectives*, edited by Adam Jaworski, Nikolas Coupland and Dariusz Galasiński, 15-51. Berlin: Mouton de Gruyter.
- Croft, William. 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Culpeper, Jonathan. 2009. "The metalanguage of IMPOLITENESS: using Sketch Engine to explore the *Oxford English Corpus*." In *Contemporary Corpus Linguistics*, edited by Paul Baker, 66-88. London: Continuum.
- Culpeper, Jonathan. 2011. *Impoliteness: Using Language to Cause Offence*. Cambridge: Cambridge University Press.
- Culpeper, Jonathan, and Michael Haugh. 2014. *Pragmatics and the English Language*. Houndmills, Basingstoke: Palgrave Macmillan.
- Culpeper, Jonathan, and Michael Haugh. (Im)politeness and sociopragmatics: In *Cambridge Handbook of Sociopragmatics*, edited by Michael Haugh, Daniel K. Kádár and Marina Terkourafi. Cambridge: Cambridge University Press.
- Dewaele, Jean-Marc. 2019. "Linguistic taboos in a second or foreign language." In *The Oxford Handbook of Taboo Words and Language*, edited by Keith Allan, 218-232. Oxford: Oxford University Press.
- Eelen, Gino. 2001. *A Critique of Politeness Theories*. Manchester: St. Jerome.
- Firth, John R. 1957. A synopsis of linguistic theory 1930-1955. In *Studies in Linguistic Analysis*, edited by John R. Firth, 1-32. Oxford: Philological Society.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press: Chicago.
- Haugh, Michael. 2012. The first-second order distinction in face and politeness research. *Journal of Politeness Research* 8(1): 111-134.
- Haugh, Michael. 2015. "Impoliteness and taking offence in initial interactions." *Journal of Pragmatics* 86:36-42.
- Haugh, Michael. 2016. "The role of English as a scientific metalanguage for research in pragmatics: Reflections on the metapragmatics of 'politeness' in Japanese." *East Asian Pragmatics* 1 (1):39-71.
- Haugh, Michael. 2018. "Corpus-based metapragmatics." In *Methods in Pragmatics*, edited by Andreas H Jucker, Klaus P Schneider and Wolfram Bublitz, 619-643. Berlin: Mouton de Gruyter.
- Haugh, Michael. 2019. "The metapragmatics of *consideration* in (Australian and New Zealand) English." In *From Speech Acts to Lay Understandings of Politeness*, edited by Eva Ogiermann and Pilar G Blitvich, 201-225. Cambridge: Cambridge University Press.

- Haugh, Michael, and Valeria Sinkeviciute. 2019. "Offence and conflict talk." In *Handbook of Language in Conflict*, edited by Jim O'Driscoll and Lesley Jeffries, 196-214. London: Routledge.
- Hübler, Alex. 2011. "Metapragmatics." In *Foundations of Pragmatics*, edited by Wolfram Bublitz and Neal R Norrick, 107-136. Berlin: Mouton de Gruyter.
- Hübler, Axel, and Ulrich Busse. 2012. "Introduction." In *Investigations into the Meta-Communicative Lexicon in English*, edited by Ulrich Busse and Axel Hübler, 1-16. Amsterdam: John Benjamins.
- Jakobson, Roman. 1960. "Linguistics and poetics." In *Style in Language*, edited by Thomas Sebeok, 350-377. Cambridge, MA: MIT Press.
- Janicki, Karol. 1999. *Against Essentialism: Toward Language Awareness*. Munich: Lincom Europa.
- Janicki, Karol. 2006. *Language Misconceived. Arguing for Applied Cognitive Sociolinguistics*. Mahwah, NJ: Lawrence Erlbaum.
- Janicki, Karol. 2010. "Lay people's language problems." *International Journal of Applied Linguistics* 20 (1):73-94.
- Janicki, Karol. 2017. "What is conflict? What is aggression? Are these challenging questions?" *Journal of Language Aggression and Conflict* 5 (1):156-166.
- Jay, Timothy. 2019. "The psychology of expressing and interpreting linguistic taboos." In *The Oxford Handbook of Taboo Words and Language*, edited by Keith Allan, 77-95. Oxford: Oxford University Press.
- Jucker, Andreas. in press 2020. *Politeness in the History of English. From the Middle Ages to the Present Day*. Cambridge: Cambridge University Press.
- Kádár, Dániel Z, and Michael Haugh. 2013. *Understanding Politeness*. Cambridge: Cambridge University Press.
- Kilgarriff, Adam, Pavel Rychly, Pavel Smrz and David Tugwell. 2004. Sketch Engine. *Proceedings of Euralex*, 105-116. Lorient, France.
- Popper, Karl. 1945. *The Open Society and Its Enemies*. London: Routledge.
- Popper, Karl R. 1979. *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press.
- Rychlý, Pavel. 2008. A lexicographer-friendly association score. In *Proceedings of Second Workshop on Recent Advances in Slavonic Natural Languages Processing, RASLAN 2008*, edited by P. Sojka and A. Horák. Brno: Masaryk University.
- Rosch, Eleanor. 1975. Cognitive Representations of Semantic Categories. *Journal of Experimental Psychology: General* 104(3): 192-233.
- Scheff, Thomas J. 2006. "Concepts and concept formation: Goffman and beyond." *Qualitative Sociology Review* 2 (3):48-64.
- Schönefeld, Doris. 1999. Corpus linguistics and cognitivism. *International Journal of Corpus Linguistics* 4:137-71.
- Silverstein, Michael. 2003. "Indexical order and the dialectics of sociolinguistic life." *Language and Communication* 23 (3/4):193-229.
- Sinclair, John. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Sperber, Dan, and Deirdre Wilson. 1986. *Relevance. Communication and Cognition*. Oxford: Basil Blackwell.
- Taylor, Charlotte. 2015. "Beyond sarcasm: the metalanguage and structures of mock politeness." *Journal of Pragmatics* 87:127-141.
- Taylor, Charlotte. 2016. *Mock Politeness in English and Italian: A Corpus-Assisted Metalanguage Analysis*. Amsterdam: John Benjamins.
- Taylor, Charlotte. 2017. "The relationship between irony and sarcasm: insights from a first-order metalanguage investigation." *Journal of Politeness Research* 12 (1):209-241.

- Terkourafi, Marina. 2005. "Beyond the micro-level in politeness research." *Journal of Politeness Research* 1 (2):237-262.
- Terkourafi, Marina. 2011. "From politeness1 to politeness2: Tracking norms of im/politeness across time and space." *Journal of Politeness Research* 7 (2):159-185.
- Trier, Jost. 1931. *Der deutsche Wortschatz im Sinnbezirk des Verstandes: Die Geschichte eines sprachlichen Feldes: Vol. 1. Von den Anfängen bis zum Beginn des 13. Jahrhunderts*. Heidelberg: Winter.
- Ullmann, Stephen. 1962. *Semantics: An Introduction to the Science of Meaning*. Oxford: Basil Blackwell.
- Verschueren, Jef. 2000. "Notes on the role of metapragmatic awareness in language use." *Pragmatics* 10 (4):439-456.
- Waters, Sophia. 2012. "'It's rude to VP": The cultural semantics of rudeness." *Journal of Pragmatics* 44:1051-1062.
- Watts, Richard. 2003. *Politeness*. Cambridge: Cambridge University Press.
- Watts, Richard, Sachiko Ide, and Konrad Ehlich. 1992. "Introduction." In *Politeness in Language. Studies in its History, Theory and Practice*, edited by Richard Watts, Sachiko Ide and Konrad Ehlich, 1-17. Berlin: Mouton de Gruyter.
- Wierzbicka, Anna. 1972. *Semantic Primitives*. Frankfurt: Athenäum.
- Wierzbicka, Anna. 1988. *The Semantics of Grammar*. (Studies in Language Companion Series 8). Amsterdam & Philadelphia: John Benjamins.
- Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. Translated by G. E. M Anscombe. 3rd ed. New York: Macmillan.

Biotnotes

Jonathan Culpeper is Professor of English Language and Linguistics in the Department of Linguistics and English Language at Lancaster University, UK. His work spans pragmatics, the history of English and stylistics, and is often underpinned by corpus methods.

Michael Haugh is Professor of Linguistics in the School of Languages and Cultures at the University of Queensland, Australia. His research interests include pragmatics, conversation analysis, intercultural communication and humour studies, with a particular focus on the role of language in social interaction.