Corpus-pragmatic perspectives on the contemporary weakening of *fuck*: The case of teenage British English conversation

Robbie Love [a,*], Anna-Brita Stenstrom [b]

[a] *Aston University, UK*
[b] *University of Bergen, Norway*

A B S T R A C T

This study examines the pragmatic functions of *fuck* among British English teenagers in casual conversation in two youth language corpora from the 1990s and 2010s. It applies a corpus-pragmatics approach to explore how the ongoing weakening of the taboo strength of *fuck* in the perception of young speakers is realised in usage data. The major functions observed involve a predominance of idiomatic, emphatic and emotionally expressive functions. Conversely, usage associated with potentially abusive functions, including literal reference to sexual intercourse, is infrequent. Our observations are interpreted in the context of delexicalization and related long-term diachronic processes, whereby contemporary usage of *fuck* among teenagers is characterised in terms of semi-delexicalized, pragmatically strengthened usage with weakened taboo status. The article also evaluates the interpretation of idiomatic usage from a functional perspective, and contributes to methodological considerations of the use of spoken corpora for pragmatic research.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Swearing by *fuck* (henceforth FUCK)[1] is spreading, not only in Britain but worldwide, and maybe not least among teenagers. As a member of what Mohr (2013) calls the "Big Six" (with CUNT, COCK, ASS, SHIT and PISS), FUCK has been characterised as being among "the worst words in English" (Mohr, 2013: 17). FUCK is of Indo-European origin (originally meaning 'to strike', Daly et al., 2004), and has cognates in related languages including the German word *ficken* ('strike' or 'hit'), the Dutch word *fokken* ('breed') and the Swedish dialectal word *focka* ('strike' or 'copulate'). In English, FUCK can be traced as far back as the 16th century and the Scottish poet William Dunbar (1460–1520), who used it in so-called 'flytings', a type of word duel intended to diss the opponent by using insulting expressions, the modern counter-part of which is 'ritual insults', which occur in contemporary teenage talk in the form of 'swearing by mother' (Drange et al., 2014: 25–59).

The contemporary spread of FUCK has been described as "a success story of almost unlikely proportions" (Ljung, 2011: 71). Now that swearing by FUCK is spreading all over the world, it is gradually losing its swearword status. This is emphasized, for

---

[1] We adopt the convention of using FUCK to refer to all morphological variants of this word. We use *fuck* only when referring to the specific morphological form (e.g. *fuck* as a singular noun).

instance, by Zenner et al. (2017: 115), who argue, with reference to the occurrence of SHIT, DAMN and FUCK in French, that FUCK is losing its taboo value — regardless of its taboo word status in the home country. In a similar vein, Jaffe (2017:103), with reference to the use of FUCK in France and Québec, states that it "arguably belongs to a repertoire of global English that belongs to everyone and no one and serves as the vehicle for stances that have an oppositional dimension". And the adoption of FUCK abroad goes even further. When studying the use of FUCK in Swedish media, Beers Fägersten (2017) noticed that it is used as a standard Swedish word. Examples such as *Han vet inte vem han fuckar med* ('He doesn't know who he is fucking with'), *Unga jurister fuckar upp språket* ('Young legal advisors fuck up the language') and *Höll några tal, sen fuckade han off* ('Held some speeches, then he fucked off') are all used in accordance with Swedish grammar (2017: 79–80). Furthermore, in the context of swearing on social media site *MySpace*, it was argued by Thelwall (2008), that FUCK should no longer be characterized as 'strong' swearing.

According to recent corpus research, swearing is most common among adolescent and young adult speakers, and FUCK is the most commonly-used swear word in contemporary casual conversation (Love, 2021). Perhaps a connected observation is that public perception of the strength of FUCK appears to have waned in recent decades, as investigated in the context of broadcast media regulation. A survey of the British public conducted by Millwood-Hargrave (2000) found that 93% of respondents considered FUCK to be 'very severe' (71%) or 'fairly severe' (22%). Twenty years later, a study by Ipsos MORI (2021) found that, relative to a large set of swear words, FUCK is (still) generally "perceived as highly offensive" (p. 18). However, a "diversity in views" was observed, whereby FUCK was "categorised variously as strong, moderate and mild by different groups" of respondents (p. 20), with the most liberal views held by younger participants (aged 18–24). So, it appears that — among younger speakers — there is less of a unified consensus on the strength of FUCK that is associated with this group being relatively more accepting of its use (see also Beers Fägersten, 2007).

Given that mild swear words tend to occur more frequently than the most offensive ones (Jay, 2009a), the high rate of usage of FUCK, while supposedly being a strongly offensive swear word, could be considered an example of what Beers Fägersten (2007: 16) calls the *swearing paradox*; namely "how this highly offensive behaviour (according to ratings studies) can also enjoy such a high rate of occurrence (according to frequency studies)". However, given that contemporary usage of FUCK (among other swear words) is skewed towards younger speakers (Love, 2021; Drummond, 2020), and younger people are most likely to consider FUCK to be relatively mild, it appears to be the case that FUCK is so common nowadays — at least among adolescents and young adults — precisely because it is better-tolerated among those groups, and therefore perceived to be milder. Based on this premise, our study focusses on the use of FUCK among younger speakers and explores the extent to which the apparently increasing acceptance of FUCK is associated with changes in the usage of FUCK — its forms and functions — that reflect its weakening status. Our observations of FUCK in our data are interpreted through the lens of usage-based theories of change, notably delexicalization, which is "the reduction of the independent lexical content of a word […] so that it comes to fulfil a particular function but has no meaning apart from this to contribute to the phrase in which it occurs" (Partington, 1993: 183). Delexicalization is part of the broader process of grammaticalization, "the dynamic, unidirectional historical process whereby lexical items in the course of time acquire a new status as grammatical, morphosyntactic forms" (Traugott, 1988: 406). In this paper, we argue that delexicalization may provide explanatory power for the apparent weakening taboo status of FUCK.

2. Swearing and pragmatics

In this study, we adopt the view of Stapleton et al. (2022: 1) that defines swearing as "the use of specific, negatively charged and often emotionally loaded terms, which are taboo in a given language/culture and thus have strong potential to cause offence". While this perspective of course recognises the potentially abusive or offensive functions of swearing, it also accounts for what is known about the many non-offensive functions of swearing. This has been observed, for example, in the study of casual conversation. Across several studies, summarised by Jay and Janschewitz (2008: 265), it has been demonstrated that "most instances of swearing are conversational; they are not highly emotional, confrontational, rude, or aggressive". And it is known that, compared to other registers, swearing is relatively common in casual conversation (Jay, 1992); corpus research (e.g. McEnery, 2006; Love, 2021) has shown that swearing is part of everyday use, and it has been found that swearing is such a common conversational feature that it does not, in this context, cause obvious harm (Jay 2009a, 2009b; Jay and Janschewitz, 2009).

Beyond abuse, swearing is used to perform a range of interpersonal functions in casual conversation, and in recent decades, linguists have sought to identify and categorise those functions. Marsden (2009) identifies four swearing categories: 'social' swearing, which strengthens group affinity; 'emphatic' swearing, which is used for emphasis; 'humorous' swearing, which occurs in joke telling; and 'aggressive' swearing, which is used to insult, threaten and ridicule. Wajnryb (2004: 26) lists three basic functions of swearing: 'cathartic', 'abusive' and 'social'. *Cathartic* swearing is an emotional outburst, i.e. when angry or in pain; *abusive* swearing is directed at a specific target (for instance a person or thing); and *social* swearing — of most relevance to our study — fulfils the purpose of joking and strengthening in-group social bonds. Andersson and Trudgill (1990: 61) also base their (non-mutually exclusive) categories around pragmatic function.

According to Ljung (2007) "[t]here are good reasons to include interjections involving swearing such as *Shit!* and *fuck* among the pragmatic markers, since "[l]ike these they may be used to express speaker attitudes, to signal the organization of text and to deliver interactional signals of various kinds". These characteristics of swearing are accentuated by Jay and

Janschewitz (2008), who argue that "[a] core use of swearing is to manage social relations and/or social interaction" with special emphasis on group membership, intimacy and trust.

Like pragmatic markers, swear words can be used to avoid silence, that is act as phatic devices. Mateo and Yus (2013) distinguish between 'praise-centred' taboo words, with no insulting effect, and 'interaction-centred' taboo words, which reinforce the social bonds between the speakers, distinguished by contextual features and tone of voice. The strong bonding, rapport-creating effect strengthens the feeling of camaraderie in teenage talk (Stenström, 2014). Pragmatic markers, like swear words, operate on three levels of conversation: the interactional, interpersonal and textual levels. Following Brinton (1996: 6), they have the following global functions.

- to mark boundaries, such as to initiate, end a discourse, or effect a topic shift
- to assist turn-taking
- to express speaker attitude
- to achieve intimacy between speaker and addressee

Turning specifically to FUCK, this swear word has a wide range of functions, including expletive (e.g. *oh fuck*), cursing (e.g. *fuck that*) and intensifier (e.g. *fucking marvellous*) (McEnery and Xiao, 2004). These are associated with meanings that have, over time, emerged from its original sense ('to strike', Daly et al., 2004), including (literally) 'to have intercourse' and (figuratively) 'to ruin' (Daly et al., 2004: 949). In terms of Politeness Theory (Brown and Levinson, 1987), even usage of FUCK that may appear on the surface to be a form of personal insult (e.g. *you fuck*) have been shown to contribute to positive politeness strategies. Daly et al. (2004) studied a small corpus of workplace conversations among factory workers in New Zealand, finding several examples (e.g. *fuck it fuck you go get your fucking legs out here*) that did not appear to convey negative affect but rather indicate solidarity:

> We suggest that, paradoxically, team members convert the negative affect and strength associated with forms of *fuck* in standard contexts into a positive attribute in its use in interaction between members of their own community of practice. The inherent strength of the canonical expletive *fuck* thus contributes to its impact when used between friends and co-workers. It is as if they are saying "I know you so well I can be this rude to you".
>
> (Daly et al., 2004: 960)

Thus, an utterance such as *You fucking bastard!* could be equivalent to *You are a great guy,* so that what looks like an insult has a purely rapport-building function, which is a reflection of 'phatic' use. In other words, the use of taboo words facilitates communication by creating and maintaining contact between the speakers, while strengthening the feeling of camaraderie, which indicates that the phatic and interactional functions go hand in hand (Mateo and Yus, 2013; Stenström, 2006, 2014).

In the context of youth language, the way swear words like FUCK are used within a group certainly reinforces social bonds and facilitates communication among the speakers within a teenage group (Stenström, 2006, 2014). For example, a type of swearing that is said to be typical of youth language is so-called 'name-calling', which is expressed in the form of sexual taboo words such as *motherfucker* and *son of a bitch*, or by insults in the form of 'swearing by mother' expressions, which involve offending someone by way of his/her mother. A distinction is made between ritual insults (*your mummy's got no lips*), name-calling (*you motherfucker*), expletive interjections (*motherfucker!*) and intensifiers (*a motherfucking bastard*). Hasund et al. (2014) argue that the pragmatic use of ritual insults realised by swearing strengthens group affinity. In a similar vein, Drange et al. (2014) discuss the pragmatic use of ritual insults realised by swearing: emphatic swearing, which is used for emphasis; humorous swearing, which occurs in joke telling; and aggressive swearing, which is used to insult (cf. Ljung, 2011).

Research in the area of swearing and pragmatics is summarised by Stapleton et al. (2022: 8), who review what is known about the positive interpersonal effects of swearing: the management of social relations and interactions; the use of swearing for rhetorical and stylistic purposes; and the use of swearing as an identity resource. For the purpose of this study, conversational swearing is seen as a means of expressing emotion, humour and verbal emphasis, and reinforcing social bonding and solidarity (Stapleton, 2010; Mateo and Yus, 2013).

3. Method

3.1. Rationale

We have shown that there has been great interest in the forms and pragmatic functions of swear words, including FUCK, from linguists working both in qualitative and quantitative disciplines. From an empirical perspective, much of the large-scale, corpus research into teenage swearing is based on data gathered in the 1990s and early 2000s. Thus, our aim in this study is to investigate how teenage use of FUCK may have changed since then by analysing data gathered more recently.

Building upon Stenström's work on the creation of the *Bergen Corpus of London Teenage Language* (COLT; Stenström and Leiv, 1993) and subsequent analysis of teenage swearing therein (Stenström et al., 2002), we are interested in the use of FUCK among teenagers in and around London, and thus sought to identify a contemporary corpus that is broadly comparable to COLT. Since no such corpus exists in isolation, we turned to a larger and more general contemporary corpus of casual

conversation — the *Spoken British National Corpus 2014* (*BNC2014*; Love et al., 2017; Stenström, 2017) — with the aim of isolating a sub-corpus of London teenage speech that would allow us to compare the use of FUCK between the 1990s and 2010s. In doing so, we aim to explore how the forms and functions of FUCK may have changed among London teenagers in casual conversation between the 1990s and 2010s. As such, the aim of this study is not to measure the strength of FUCK in the corpus data. Instead, we approach the data with the assumption that, in recent decades, there has been a reduction in the taboo strength of FUCK among younger speakers, and we seek to establish what evidence (if any) the usage data may provide that informs understanding of the apparent weakening of FUCK in the perception of younger speakers.

Therefore, this study is based upon the diachronic comparison of two language samples separated by little over two decades. Although language change is typically studied over longer periods of time (e.g. D'Arcy, 2015), our approach is informed by that of Aijmer (2022: 9), who defends the use of "comparable present-day corpora" as a means with which to describe "[O]n-going changes in a language […] in an emergent and usage-based perspective where meanings are negotiated by speakers in real time". As such, while we acknowledge that, in terms of ongoing processes of change, our findings should be subjected to further enquiry with additional, comparable datasets, our approach responds to the "lack of research using empirical data to study on-going changes in pragmatics and discourse" by exploring contemporary change that is "motivated by social factors and continuously evolving discourse practices and sociocultural norms" (Aijmer, 2022: 9).

3.2. Data: COLT

To facilitate the construction of COLT, 31 boys and girls were recruited in 1992 to record conversations with friends of the same age in various out-of-school situations, with no grown-up interference. This resulted in a corpus comprising transcripts of conversations from a total of 83 teenage speakers. The recordings resulted in transcripts of varying lengths (the shortest three conversations, the longest 39), totalling 614,506 tokens. All the speakers came from London districts, notably Barnet, Brent Camden, Enfield, Hackney, Hertfordshire, Islington, Richmond, Tower Hamlets and Westminster. The number of speakers per conversation varies from two to five, the majority with both male and female speakers.[2] The COLT conversations were recorded in various settings, for instance, in a school yard, in a street, in a park, at a cafe, at home, etc.

When preparing to analyse the COLT data, we noted the existence of two 'versions' of the corpus. The original version of COLT was transcribed by the British National Corpus research team and included in the BNC1994 (BNC Consortium, 2007). However, the COLT research team at Bergen later checked and made corrections to the original transcripts (in doing so, increasing the size of the corpus by "at least 15 per cent")[3] and made the new set of transcripts available as a standalone corpus. Since our aim was to compare the COLT data to transcripts from the new BNC2014, we initially hoped that, for ease of comparability between sub-corpora of the BNC1994 and BNC2014, and for ease of access to speaker metadata, we would be able to isolate the texts in the BNC1994 that had been originally 'donated' from the COLT project. Upon investigation, we found that this was not possible, because (a) there does not appear to be a record of which BNC1994 texts came from COLT and (b) in many cases, BNC1994 'texts' contain transcripts of multiple conversations, and we observed instances where COLT and non-COLT transcripts were contained within the same BNC1994 'text'. With the possibility of isolating the BNC version of COLT eliminated, we proceeded with the 'corrected' version of COLT with the caveat that, due to the formatting of these files, access to speaker metadata would be limited; as reported by Torgersen et al. (2011: 99): "COLT contains the speech of the friends, family and teachers of the recruits […], with no consistent encoding of their age, sex, ethnicity and residence". COLT was downloaded from the CLARINO UiB Portal[4] and then uploaded to Sketch Engine (Kilgarriff et al., 2014) for analysis.

3.3. Data: Spoken BNC2014 sub-corpus

As mentioned, we sought to derive a sub-corpus from the Spoken BNC2014 that is broadly comparable to COLT for the purpose of analysing the usage of FUCK in teenage casual conversation. The Spoken BNC2014 is an 11-million-token corpus of casual conversation recorded among members of the UK public in the period 2012–2016 (Love, 2020). While the linguistic register of the Spoken BNC2014 is comparable to COLT in that they both comprise casual conversations, we noted differences in the procedures used to compile COLT and the Spoken BNC2014 that factored into our approach. Firstly, the compilation of COLT is an example of a 'specialised' corpus design (see e.g. Koester, 2022) in that its participants were recruited based on a narrow set of sampling criteria — namely teenagers living in London. On the other hand, the Spoken BNC2014 is a 'general' corpus that represents a variety of a language (see e.g. Clancy, 2022) — in this case a broad sample of British English speakers of all ages, mostly across England. Secondly, as per the BNC1994 recording procedure (Crowdy, 1993), the COLT conversations were recorded surreptitiously, while the Spoken BNC2014 conversations were recorded with the prior knowledge and

---

consent of all participants. This difference may have some impact on the amount of swearing observed in the Spoken BNC2014; there is relatively less swearing in the Spoken BNC2014 compared to the Spoken BNC1994 (Love, 2021). The use of swearwords in the Spoken BNC2014 might have been hampered by the speakers being aware that they were being recorded. That said, there is evidence that many COLT participants were probably aware of being recorded; there are over 500 references to the *tape* or *taping* and over 300 references to the *recorder* or *recording* in the transcripts, as reflected in the following extract from COLT:

A: You wanna listen to this one
B: Ha? You lot ain't supposed to know I'm **taping** (COLT, B132607)

Consequently, while the difference in speaker awareness of being recorded is difficult to estimate, the broader ethical procedures do differ, and this difference is unavoidable with respect to developments in the ethics of academic research in the intervening years that would preclude the secret recording of conversations. Besides, the gathering of informed consent for the Spoken BNC2014 recordings did bring about a major benefit: a much richer and more complete set of metadata for each speaker, as this information could be gathered at the point of participants reading and signing the consent form (Love, 2020: 61).

With these issues in mind, we sought to identify texts in the Spoken BNC2014 that could be compared to those in COLT. The design criteria identified in COLT are: teenage speakers, from London, holding conversations predominantly with other teenagers (with minimal or no input from speakers of other ages). In attempting to replicate these criteria using texts from the Spoken BNC2014, we initially identified a total of 54 teenage speakers in the corpus. However, according to the metadata, only seven of these speakers were born in London. In order to increase the size of the sub-corpus, we decided to broaden the inclusion criteria to allow speakers from a larger area of south-east England. In total, we identified 15 teenage speakers from the south-east of England who participated in a total of 35 teenager-only conversations across 25 h of recordings. These texts were isolated to form a sub-corpus of 307,104 tokens, roughly half the size of COLT (see Appendix for textual metadata). The speakers are between 13 and 18 years old, all from the south-England area (Cambridge, Southall, Staines, Slough, High Wycombe, Camberley, Ashville, Aldershot, Frimley, Winchester, and Falmouth). The number of speakers per conversation varies from two to five, and the typical pattern is girl–girl and boy–boy conversations – with few exceptions. Most of these recordings took place in a 'wider family circle' and the topics centred on school, food and drink, TV series, teachers, friends, and university courses (see Appendix for speaker metadata). The Spoken BNC2014 is already available in Sketch Engine,[5] so we were able to create the sub-corpus in the tool via identification of the corpus text filenames. Table 1 presents a summary of the corpora used in the study.

There are limitations to our approach. In lieu of compiling a new, specialised corpus of London teenage speech to match in the design of COLT, we have attempted to 'retrofit' an existing corpus (the Spoken BNC2014) by extracting, as a sub-corpus, an approximation of the COLT sampling frame. This exercise necessarily involved compromise and resulted in a sub-corpus (see Appendix) that is not a 'perfect match' with COLT. Nonetheless, it does share major features that we judged deemed worthy of exploration and comparison to COLT: casual conversational data, recorded among teenage L1 speakers of British English from the south east of England (London and surrounding areas), and transcribed orthographically by a research team who intended to facilitate comparison between the original BNC1994 (which contains COLT) and the BNC2014.

Table 1
Information about the corpora used in the study.

	COLT	Spoken BNC2014 sub-corpus
Number of speakers	83	15
Number of texts	377	35
Number of tokens	614,506	307,104
Available speaker metadata	None	Many categories including gender, age, socio-economic status

3.4. Analytical approach

Within Sketch Engine, we first used the Corpus Query Language[6] search query *[word=".*fuck.*"]* to search for all morphological forms of ꜰᴜᴄᴋ in both corpora; the wildcard character (*) captures any number of unspecified characters that may be affixed, in this case, at the beginning and/or at the end of the base form ꜰᴜᴄᴋ. All instances of ꜰᴜᴄᴋ in both corpora were extracted, along with their immediate co-text, to produce two spreadsheets comprising the concordance lines of ꜰᴜᴄᴋ from both corpora. We then annotated all instances manually via close inspection of the concordance lines, which provide

---

sufficient linguistic context to classify the usage of FUCK. Each instance was categorised qualitatively according to major word class and swearing function.

Table 2 presents our word class coding scheme. We coded the major word classes of noun, verb, adjective and adverb, to which we added the category of interjection, which refers to "exclamative utterances" that "do not easily fit into the major word classes" (Carter and McCarthy, 2006: 224). The inclusion of this category was necessary as it is well-established that, in interpersonal terms, one of the many purposes of swearing is to express emotion via expletives (see e.g. Stapleton, 2010), which may be syntactically independent and therefore categorised separately to the major word classes.

Table 2
Coding scheme for major word class.

Category	Criteria (Carter and McCarthy, 2006)	Example(s)
Adjective	"describe features and qualities of entities (people, animals and things) denoted by nouns or pronouns" (p. 439)	*this chair is fucked; in the fucking car*
Adverb	"typically modify verb phrases, adjectives and other adverbs. Some adverbs modify whole clauses or sentences" (p. 453)	*fucking marvellous*
Noun	"referring expressions […] used to refer to particular instances or general classes of people and things" (p. 318)	*you stupid fuck; fat as fuck*
Verb	"Lexical verbs can stand alone. Lexical verbs have meanings denoting actions, events and states, and belong to an open class" (p. 419)	*he fucked off; fuck you*
Interjection	"Exclamative utterances […] that do not easily fit into the major word classes (noun, verb, adjective, adverb)" (p. 224)	*(oh) fuck*
Unknown	It is not possible to identify the word class due to insufficient context.	N/A

Turning to function, our coding scheme (Table 3) was developed from the functional categorisation scheme adopted by McEnery and Xiao (2004: 257-8) in their study of FUCK in the original British National Corpus. Based on the 'categories of insult' scheme developed by McEnery et al. (1999, 2000) to capture the major functions of English swear words (see also McEnery, 2006), it was adapted by McEnery and Xiao's (2004) to remove the categories that "do not apply to *fuck*" (p. 257). The scheme provides a formal basis with which to observe − at the microstructural level − the quantitative distribution of the interpersonal functions of FUCK, and has been shown to be informative in research in corpus pragmatics (McEnery et al., 2023).

Table 3
Coding scheme for swearing function.

Category	Criteria	Example(s)
G	General expletive	*(oh) fuck*
P	Personal insult referring to defined entity	*you fuck/that fuck*
C	Cursing expletive	*fuck you/me/him/it*
D	Destinational usage	*fuck off/he fucked off*
L	Literal usage denoting taboo referent	*he fucked her*
E	Emphatic intensifier	*fucking marvellous/in the fucking car*
O	'Pronominal' form	*like fuck/fat as fuck*
I	Idiomatic 'set phrase'	*fuck all/give a fuck/thank fuck*
A	Predicative negative adjective	*this game is fucked*
X	Metalinguistic or unclassifiable due to insufficient context	The use of the word "fuck"/*you never fucking*

As discussed (in the context of corpus-based research) by Larsson et al. (2020), inherent in manual qualitative annotation is the potential for inconsistency in coding which can be caused by a variety of factors, including ambiguity in the coding scheme and random human error. Before finalising our coding schemes, we piloted them on a sample of concordance lines. We also asked another linguist (otherwise not involved in the study) to code the same sample after receiving instruction on the use of the schemes. We measured inter-rater agreement/reliability on a 95% confidence ($+/−5\%$) random sample (Israel 1996; Moinester and Gottfried 2014)[7] of 270 concordance lines (an equal sample of 135 examples from each corpus). We calculated inter-rater agreement using Cohen's kappa (κ), a chance-adjusted measure of agreement (Cohen, 1960). The mean rate of agreement between our coding and that of our colleague was, for word class, 82.2% ($\kappa = 0.79$), and for swearing function 84.4% ($\kappa = 0.83$), indicating 'substantial' agreement for word class and 'almost perfect' agreement for swearing function (Landis and Koch, 1977: 165).

While this provided evidence that our coding could be reproduced with a high level of reliability, we sought nonetheless to investigate potential causes for the relatively few cases where disagreement occurred. For word class, this resulted in clarification that emphatic *fucking*, when used to modify nouns and pronouns (e.g. *I have so many leftover fucking limes*), would be classified as an adjective as opposed to an adverb. For swearing function, one category that McEnery and Xiao (2004) removed from the original 'categories of insult' scheme (McEnery et al., 1999, 2000), on the grounds that it did not apply to FUCK, is

---

[7] This was calculated using confidence sampling software: https://www.calculator.net/sample-size-calculator.html

category A: (predicative negative adjective). During our piloting, we encountered several examples of predicative negative adjective FUCK in both COLT and the BNC2014 data (e.g. *last night was completely fucked you know*; COLT, B141707), so we reintroduced this category to the scheme.

More broadly, we found patterns between swearing function and word class which led us to include guidance on typical word class to the function scheme in order to reduce ambiguity (e.g. destinational usage of FUCK is invariably performed by a verb, and emphatic intensifier FUCK occurs only as an adjective or adverb). Indeed, the relationship between swearing function and word class is acknowledged by McEnery (2006: 28), who notes that "[p]arts of speech are clearly important to the categorisation scheme, but the scheme itself is not simply a relabelling of parts of speech". However, the only function category for which it was not possible to narrow the criteria according to word class was category I (idiomatic 'set phrase'), as we observed that idiomatic usage of FUCK may (a) involve any word class, and (b) by definition, make it difficult to reliability deconstruct phrasal units into component word classes. For example, for *shut the fuck up*, there are (at least) two possible analyses: *fuck* could be classified as a noun (within the noun phrase *the fuck*), but it could also be considered adverbial, since *the fuck* is modifying the phrasal verb *shut … up*).

Furthermore, disagreement over the idiom status of some phrases involving FUCK contributed to the imperfect inter-rater agreement score. To resolve this, we sought to identify a set of examples that both coders agreed upon as idiomatic and use these to form a closed class (listed in Table 7, Section 4.3). Cases that were disagreed upon were excluded from category I and reassigned to one of the remaining categories.

Following these adjustments, the same sample of 270 concordance lines was re-coded independently, and the inter-rater agreement measures improved: the mean rate of agreement between our revised coding and that of our colleague was, for word class, 98.1% ($\kappa = 0.98$), and for swearing function 98.9% ($\kappa = 0.98$), indicating 'almost perfect' agreement for both coding schemes (Landis and Koch 1977: 165).

4. Findings

4.1. Frequency and forms of FUCK

There are significantly more instances of FUCK in COLT than the Spoken BNC2014 teenage sub-corpus. We retrieved 727 forms of FUCK in COLT (relative frequency: 652 per million tokens) and 171 forms of FUCK in the Spoken BNC2014 sub-corpus (relative frequency: 547 per million tokens). According to the log-likelihood test, the difference in relative frequency is statistically significant at $p < 0.0001$ (Log-likelihood = 90.79; Log Ratio = 1.09),[8] replicating the significant difference observed for general swearing among all speakers in the Spoken British National Corpora (Love, 2021: 11). This difference can likely be explained by the same reasons suggested for there being less swearing overall in the Spoken BNC2014 (as discussed in Section 3.3).

Table 4 shows the distribution of morphological forms of FUCK in both corpora. The headline observation here is that most swearing by FUCK is performed by three forms: *fucking*, *fuck* and *fucked* − this aligns with the findings of McEnery and Xiao (2004: 258) for the forms of FUCK across the whole of the BNC1994. Interestingly, though the three most common forms are the same for both corpora the ranking differs; *fucking* has been displaced as the most common form by FUCK in the more recent data. The possible cause of this is explored in Section 4.3 once word class and pragmatic function have been taken into account. More broadly, Table 4 shows an apparent reduction in the diversity of forms in that several rarer forms present in COLT (e.g. *motherfucker*, *unfuckingtouchable*) are not found in the Spoken BNC2014 teenage sub-corpus; this may be a product of the 2014 sample being around half the size of COLT, providing less opportunity for rarer forms to emerge.

Table 4
Morphological variants of FUCK in COLT and the Spoken BNC2014 sub-corpus.

Word form	COLT		Spoken BNC2014 sub-corpus	
	Frequency	%	Frequency	%
fucking	395	54.33	60	35.09
fuck('s)	239	32.87	88	51.46
fucked	62	8.53	20	11.70
fucker	10	1.38	0	0.00
motherfucker(s)	10	1.38	0	0.00
mother(-)fucking	5	0.69	0	0.00
fucks	5	0.69	2	1.17
unfuckingtouchable	1	0.14	0	0.00
fuckers	0	0.00	1	0.58
Total	727	100	171	100

---

[8] Log-likelihood was calculated using the UCREL *Log-likelihood and effect size calculator*: https://ucrel.lancs.ac.uk/llwizard.html.

4.2. Word class

Fig. 1 shows the proportional distribution of word class categories of all instances of FUCK in the two datasets. Statistically, the difference in distribution of word class categories between COLT and the Spoken BNC2014 teenage sub-corpus is significant (Chi-squared (5) = 32.53, $p < 0.0001$) with a small effect (Cramer's V = 0.19, 95% CI [0.11, 0.245]).[9] Three word class categories — *adjective, adverb* and *verb* — take up a lower percentage of FUCK in the BNC2014 data than in COLT. This corresponds with two word class categories — *interjection* and *noun* — taking up a higher percentage in the BNC2014 data. Of particular note is the use of FUCK as a noun, which accounted for only 10.0% of instances in COLT but accounts for 24.6% of instances in the Spoken BNC2014 teenage sub-corpus; this is the biggest difference across all the word classes.
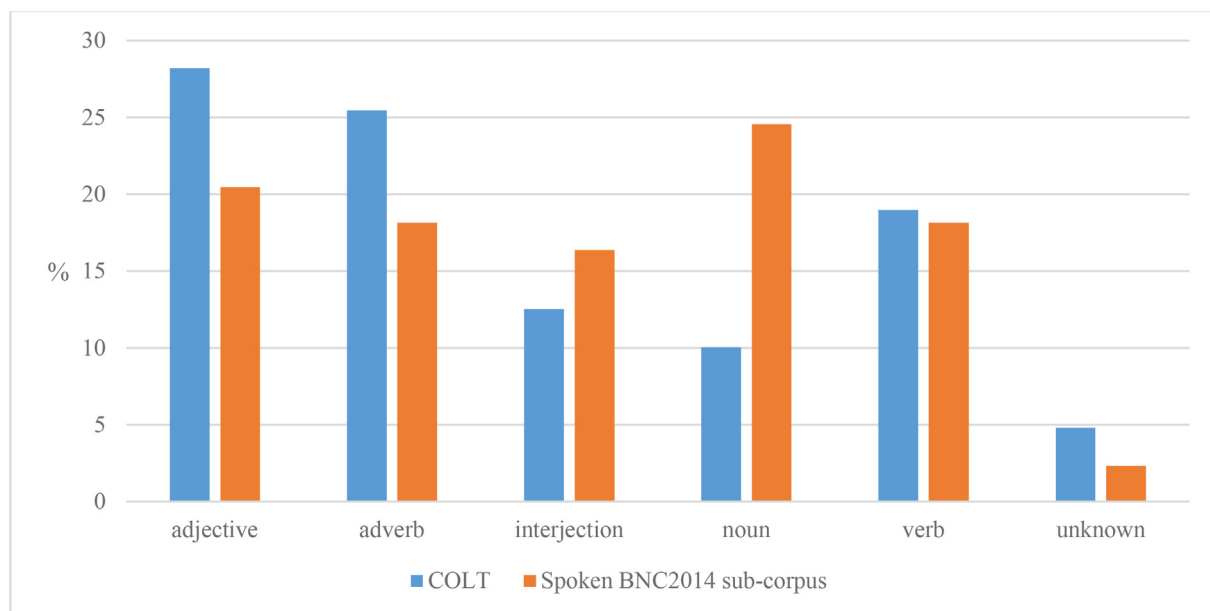

Fig. 1. Distribution (%) of major word classes of FUCK in COLT and the Spoken BNC2014 sub-corpus.

Table 5 shows the morphological forms of FUCK across the word class categories, the distribution of which is significant in both corpora.[10] This sheds some light on the apparent increase in the use of FUCK as a noun; it appears to be driven by a relative increase in the use of singular noun *fuck*; firstly, the *fuck* form accounts for 72.6% (53 out of 73) of FUCK as a noun in COLT, but 95.2% (40 out of 42) in the BNC2014 data. Secondly, nominal usage of the *fuck* form accounts for 22.2% (53 out of 239) of all FUCK usage in COLT, but 45.5% (40 out of 88) in the BNC2014 data. Together, these point towards a newfound emergence of the use of singular noun *fuck* that is explored from a functional perspective in Section 4.3.

Table 5
Distribution of word forms of FUCK across word class categories in COLT and the Spoken BNC2014 sub-corpus.

Word form	COLT							LL score	Sig. level
	Total	adjective	adverb	interjection	noun	verb	unknown		
fucking	395	157	181	19	0	12	26	831.15	*p* < 0.0001
fuck('s)	239	0	3	72	53	103	8		
fucked	62	44	0	0	0	18	0		
fucker	10	0	0	0	10	0	0		
motherfucker(s)	10	0	0	0	10	0	0		
mother(−)fucking	5	4	0	0	0	0	1		
fucks	5	0	0	0	0	5	0		
unfuckingtouchable	1	0	1	0	0	0	0		
Total	727	205	185	91	73	138	35		

[9] Chi-squared and Cramer's *V* were calculated using *Lancaster Stats Tools online* (Brezina, 2018): http://corpora.lancs.ac.uk/stats/toolbox.php.
[10] Log-likelihood was calculated using *Lancaster Stats Tools online* (Brezina, 2018): http://corpora.lancs.ac.uk/stats/toolbox.php.

Table 5 (*continued*)

Word form	COLT								
	Total	adjective	adverb	interjection	noun	verb	unknown	LL score	Sig. level
Percent	100	28.20	25.45	12.52	10.04	18.98	4.81		
Word form	Spoken BNC2014 sub-corpus								
	Total	adjective	adverb	interjection	noun	verb	unknown	LL score	Sig. level
fuck('s)	88	0	0	24	40	22	2	209.41	p < 0.0001
fucking	60	21	31	4	1	1	2		
fucked	20	14	0	0	0	6	0		
fucks	2	0	0	0	0	2	0		
fuckers	1	0	0	0	1	0	0		
Total	171	35	31	28	42	31	4		
Percent	100	20.47	18.13	16.37	24.56	18.13	2.34		

4.3. Swearing function

Table 6 shows the distribution of functional categories across FUCK in the COLT and BNC2014 data. The difference in the distribution between the corpora is statistically significant (Log-likelihood = 69.45, *p* < 0.0001).[11] Category E (emphatic intensifier) is the most common category in COLT, accounting for just under half (49.7%) of all instances of FUCK, and the second most common category in the Spoken BNC2014 teenage sub-corpus, accounting for 31.6%. This category comprises emphatic adjectives (1−2) (COLT = 22.1% of FUCK; BNC2014 = 12.3% of FUCK) and adverbs (3−4) (COLT = 25.0% of FUCK; BNC2014 = 18.1% of FUCK). Among category E cases, the proportion of adjectival and adverbial usage differs between COLT (46.9% adjective; 53.1% adverb) and the Spoken BNC2014 sub-corpus (40.4% adjective; 59.1% adverb).

(1) *[name] stuck a few pictures on his erm **fucking** calculator case* (COLT, B141906)
(2) *my teacher was a weird **fucking** psycho who fed us conspiracy theories* (BNC2014, S784)
(3) *Someone tell him about it I'm **fucking** getting fed up with telling every single person* (COLT, B133901)
(4) *I'm gonna be really **fucking** annoying about it* (BNC2014, SMRV)

Table 6
Functional categories of FUCK in the COLT and BNC2014 data.

Code	Description	COLT		Spoken BNC2014 teenage sub-corpus		% diff
		Frequency	%	Frequency	%	
A	Predicative negative adjective	29	3.99	7	4.09	0.10
C	Cursing expletive	18	2.48	8	4.68	2.20
D	Destinational usage	59	8.12	5	2.92	−5.19
E	Emphatic intensifier	361	49.66	54	31.58	−18.08
G	General expletive	70	9.63	26	15.20	5.58
I	Idiomatic 'set phrase'	104	14.31	65	38.01	23.71
L	Literal usage denoting taboo referent	40	5.50	3	1.75	−3.75
O	'Pronominal' form	0	0.00	0	0.00	0.00
P	Personal insult referring to defined entity	23	3.16	1	0.58	−2.58
X	Metalinguistic or unclassifiable due to insufficient context	23	3.16	2	1.17	−1.99
	Total	727	100	171	100	

The biggest difference between the functional categories in COLT and the Spoken BNC2014 teenage sub-corpus is that category I (idiomatic 'set phrase') is almost three times as common, proportionally, in the BNC2014 data than in COLT, making it the most populated functional category in the BNC2014 data (5−8). Upon investigation of this category, we arrive at the likely explanation for our observations about singular noun *fuck* in Sections 4.1 and 4.2: idiomatic usage of FUCK is heavily associated with nominal usage, as shown in Table 7, which presents the idiomatic structures identified in the data. This shows that the most common idiom structure in both corpora is *[WH- THE FUCK]*, whereby an interrogative pronoun or adverb (usually *what*, as in *what the fuck*, but also *who, how, where* and *when*) is post-modified by noun phrase *the fuck*. In COLT, 44 out of 104 (42.3%) instances of FUCK in category I are nominal, and in the BNC2014 data, this usage amounts to 39 out of 65 (60.0%). In both cases, noun is the most common word class category attributed to idiomatic FUCK and, as already shown, most nominal

---

[11] Due to lower-than-sufficient expected frequencies, the log-likelihood test was used instead of chi-squared (Brezina, 2018: 113). This test necessarily excludes category O ('Pronominal' form), as it has a frequency of zero in both corpora.

usage of FUCK in both corpora is performed by singular noun form *fuck* (5−6). This supports McEnery and Xiao's observation that the singular *fuck* "is most frequently used idiomatically, as in *what the fuck*, *for fuck's sake*, *give a fuck* and *fuck up/around/about*" (2004: 258-9).

(5) *we were just sitting here thinking okay **what the fuck** is going on?* (BNC2014, SXR9)

(6) *oh **for fuck's sake** when was that like four AM or something?* (BNC2014, SUZS)

(7) *did I **fuck something up**?* (BNC2014, S2LD)

(8) *Jesus that's like two Jager Bombs each which […] will **fuck you over*** (BNC2014, SXK5)

Table 7
Idiomatic FUCK in COLT and the Spoken BNC2014 sub-corpus.

COLT				Spoken BNC2014 sub-corpus			
Rank	Idiom structure	Count	%	Rank	Idiom structure	Count	%
1	[WH-] THE FUCK	28	26.92	1	[WH-] THE FUCK	25	38.46
2	FUCKING HELL	17	16.35	2	FUCK_v UP	9	13.85
3	FUCK_v UP	11	10.58	3	FUCKED_adj UP	8	12.31
4	FOR FUCK'S SAKE	7	6.73	4	FOR FUCK'S SAKE	6	9.23
5	FUCKED_adj UP	6	5.77	5	GIVE A FUCK	3	4.62
6	GIVE A FUCK	5	4.81	6	FUCK [PRON] OVER	3	4.62
7	SHUT THE FUCK UP	5	4.81	7	FUCKING HELL	2	3.08
8	CAN'T BE FUCKED	5	4.81	8	SHUT THE FUCK UP	2	3.08
9	FUCKED_adj OFF	4	3.85	9	FUCK WITH [NOUN]	2	3.08
10	FUCK WITH [NOUN]	3	2.88	10	[ADJ] AS FUCK	2	3.08
11	FUCK ALL	3	2.88	11	FUCK IT	1	1.54
12	FUCK AROUND	3	2.88	12	FUCK ABOUT	1	1.54
13	FUCK ME OFF	2	1.92	13	FUCK KNOWS	1	1.54
14	[ADJ] AS FUCK	1	0.96				
15	BEAT THE FUCK OUT OF	1	0.96				
16	FUCK A DUCK	1	0.96				
17	GET THE FUCK	1	0.96				
18	HEAD FUCK	1	0.96				
	Total	104	100		Total	65	100

The remaining categories differ far less substantially in frequency between the two corpora. Nonetheless, there are smaller differences in four of the categories − G, L, P and D − that may, when considered together, help in interpreting the bigger differences in E and I (as discussed in Section 5). Firstly, there is a higher proportion of category G (general expletive) in the BNC2014 data (the third most frequent function in both corpora, rising from 9.6% to 15.2%), accounting largely for cases where FUCK occurs in isolation without grammatically modifying another word (9).

(9) A: *no whoa whoa whoa whoa*
   B: *ah **fuck*** (BNC2014, S2LD)

Conversely, though already long-observed to be a rare usage of FUCK nowadays (e.g. McEnery and Xiao, 2004), there is a smaller proportion of category L (literal usage denoting taboo referent) falling from 5.5% to 1.8% (10−11).

(10) *I know for definitely sure that Miss's **fucked** one of the upper sixth* (COLT, B141405)

(11) *the beginning bit when she's **fucking** a man she sticks an axe through him* (COLT, B132901)

Thirdly, category P (personal insult referring to defined entity) has only one instance (12) in the BNC2014 data (account for 0.58% of functions). This compares to 3.2% in COLT, among which are several examples more typical of impolite insults, specifically personalised negative vocatives (13, 14) and third-person negative references (Culpeper, 2011: 135) (15).

(12) *Whatsapp **fuckers** on our group chat* (BNC2014, S784)

(13) *That's why I'm cleaning it you fat **fuck*** (COLT, B142002)

(14) *What you want **fucker**?* (COLT, B138102)

(15) *[name] is a really fat ugly rough **motherfucker*** (COLT, B141606)

Finally, category D (destinational usage) accounts for 8.1% of FUCK in COLT (16), falling to 2.9% in the BNC2014 data (17).

(16) *Shall I call him, shall I call him then you tell him to **fuck off**?* (COLT, B142003)

(17) *yeah I know I felt like such a dick after that I was like I was like no **fuck off** and then erm they had presents for me* (BNC2014, S29Q)

5. Discussion

5.1. Pragmatic change

In this section, we interpret our observations through the lens of delexicalization, a long-term diachronic process that can be considered a continuum between the lexical and grammatical; linguistic units can be observed at various stages of partial

delexicalization (or *semi-delexicalization*), whereby evidence of both relatively lexicalized and relatively grammaticalized usage may be observed simultaneously (Partington, 1993; Lorenz, 2002). Delexicalized/grammaticalized words eventually undergo semantic bleaching (Herda, 2017: 203), the early stages of grammaticalization are found to be associated with "strengthening, most especially strengthening of the expression of speaker involvement" (Traugott, 1988: 407). Pragmatic strengthening is initially observed through "a redistribution or shift, not a loss, of meaning" (Hopper and Traugott, 2003: 94), that "affects radically the function of an item" (Tognini-Bonelli, 2000: 231). As additional functions are gained, semi-delexicalized words "can be used in a wider range of contexts" (Heine et al., 1991: 109) and, accordingly, are used with increased frequency (Bybee and Pagliuca, 1985).

From a pragmatic perspective, this provides a breeding ground for the emergence and strengthening of figurative and interpersonal functions (expressed broadly as *pragmatic meaning*, Traugott, 1988; Traugott and König, 1991; Hopper and Traugott, 2003). This appears to be the case for the use of ꜰᴜᴄᴋ in both COLT and the Spoken BNC2014 sub-corpus; at the lexical end of the continuum, there are (albeit relatively few) literal references to sexual intercourse (category L). Towards (but not actually at) the grammatical end, there is emphatic and idiomatic ꜰᴜᴄᴋ (categories E and I respectively) — both of which are associated with delexicalization.

Starting with the emphatic function, *fucking* can be considered a "semi-delexicalised" (Partington, 1993: 188) intensifier; in a study of *fucking* in the Spoken British National Corpora, adverb *fucking* was found to modify adjectives with both positive, negative and neutral connotations, a sign of partial delexicalization (Aijmer, 2018: 70). The prevalence of emphatic *fucking* in our data reflects contemporary usage of *fucking* for social and emotive emphasis and "evaluative meaning construction" (Pan, 2022: 397). Our observation that adverbial *fucking* appears to be winning over adjectival *fucking* as the predominant word class is resemblant of subjectification (Traugott, 1988, 1989). The trajectory of classifier > intensifier has been traced in the development of intensifying meanings in adjectives (e.g. He, 2017), whereby "use of intensifiers with predicate adjectives could be taken as evidence for a later stage in the delexicalization process" (Ito and Tagliamonte, 2003: 271).

As for the idiomatic function, idiomatic ꜰᴜᴄᴋ involves conventionalised structures that are analysed above the level of the word, i.e. at the phrasal level. The substantial increase in the proportion of idiomatic ꜰᴜᴄᴋ in our data may be a sign of idiomatization, described as "the diachronic element of lexicalization" (Brinton and Traugott, 2003: 56), i.e. the trajectory towards the end-of-life of a lexical item. This represents a departure from the use of ꜰᴜᴄᴋ as an independent unit and instead "an extended unit of meaning where the function of the whole is different from the function of the individual parts" (Tognini-Bonelli, 2000: 229). Reduction in compositionality and increase in semantic opacity, as triggered by idiomatization, is associated with delexicalization (Kecskés, 2000), and a similar diachronic shift was observed in late-twentieth century written British English by McEnery and Xiao (2004), suggesting the potential for the existence of a broader trend.

We have also noted the distribution of four function categories that are not as frequent but nonetheless relevant. Category G (general expletive) is the third most frequently observed function of ꜰᴜᴄᴋ, behind idiomatic and emphatic functions. It represents exclamative usage that is pragmatically variable; not only is it associated with catharsis (*intra*personal emotional processing), but from an *inter*personal perspective these are pragmatic markers of stance (Ljung, 2011), associated with the contextually-dependent expression of a range of emotions, from anger and frustration to happiness and excitement (Stapleton, 2010: 294). In casual conversation between highly familiar members of a social group — in our case, teenagers — the frequent use of ꜰᴜᴄᴋ with expletive function can be seen as a means of affiliating with the group (Stenström, 2006; Stapleton et al., 2022), signalling positive affect (Daly et al., 2004) and distinguishing their teenage identity from that of older adults (Aijmer, 2018). From this perspective, expletive usage can be interpreted similarly to emphatic usage; expletive FUCK is semantically versatile, functioning to amplify the speaker's expression of their strength of feeling in a variety of ways, positively and negatively. In other words, expletive usage of FUCK is generalized; as a result of delexicalization, its "meanings are emptied of their specificities" (Bybee and Pagliuca, 1985: 63), but the residual taboo status of FUCK allows it, at least for now, to be used for pragmatic strength (Traugott, 1988).

The other three categories of note are L (literal usage denoting taboo referent), P (personal insult referring to defined entity) and D (destinational usage). These are the remaining function categories to have decreased in usage (proportionally) from one corpus to the next (combined total: 16.78% in COLT; 5.26% in Spoken BNC2014 sub-corpus). In terms of contemporary usage of ꜰᴜᴄᴋ, literal reference to sexual intercourse is the most strongly lexicalized function and — of course — the source of the taboo status of the swear word. In the context of (im)politeness, category P includes personalised negative vocatives and third-person negative references, and category D is a dismissal (Culpeper, 2011: 135). These are "conventionalized impolite formulae" (Culpeper, 2011: 133) that involve explicitly defining the target of the insult, which is potentially face-threatening. Arguably, L, P and D are among the categories most likely to involve 'abusive' swearing, which is indicative of strong taboo status (Wajnryb, 2004). However, these categories occur very infrequently, relative to the semi-delexicalized functions (especially so in the Spoken BNC2014 sub-corpus). Even among those few that remain, it is likely that there are instances of mock impoliteness (Culpeper, 2011; cf. Daly et al., 2004); for example, drawing on findings from a multi-billion-word corpus of English, it is estimated that at least 12% of instances of the impoliteness formula YOU + NP (e.g. *you fat fuck*) are "not intended or perceived as impolite in the settings in which they are used" (Van Olmen and Marta Andersson, 2023: 34).

5.2. Summary

Our study has described the use of ꜰᴜᴄᴋ among two contemporary samples of British English teenage speakers, separated by little over two decades. Overall, a picture emerges of contemporary ꜰᴜᴄᴋ that is highly emphatic, idiomatic, figurative and emotionally expressive, and very infrequently literal or otherwise (potentially) abusive. This functional distribution can be explained through delexicalization, which, within the early stages of grammaticalization, predicts pragmatic strengthening (Traugott, 1988). Based on our analysis, it appears to be the case that contemporary ꜰᴜᴄᴋ is delexicalized sufficiently to have undergone pragmatic strengthening and (partial) idiomatization.

Now we return to the rationale for this study: the question of the extent to which the apparent weakening of the taboo status of ꜰᴜᴄᴋ among young speakers — as observed in perceptual studies (Millwood-Hargrave, 2000; Ipsos MORI, 2021) may be evident in corpus data. Delexicalization arguably provides explanatory power in accounting for long-term diachronic changes in the strength and functions of swear words. Lafreniere et al. (2022: 909) draw upon delexicalization to describe the gradual loss of "original (taboo) meanings", such that swear words become "mixed-meaning expressions" that convey both topic meaning (figuratively extended from the denotative meaning) and interpersonal meaning (strength of feeling). The differences between our earlier (1990s) and later (2010s) samples suggest, at least on the surface, an incremental step further in this direction; proportionally, there is even less literal usage, indicative of "increasing abstraction: from physical place terms […] to discourse terms" (Stubbs, 2002: 14), and there is more of the figurative usage that is likely to perform positive as opposed to negative (e.g. abusive) interpersonal functions. Put together, this pattern of usage would seem to accord with an interpretation of weakened taboo status, and thus we would conclude that the weakening of FUCK among teenage speakers is indeed 'visible' in usage data in the ways we have described.

So there are principles, derived from theories of diachronic processes associated with grammaticalization, which may help to explain our observations. However, as discussed in Section 3.1, we are careful not to assert with certainty that the differences we have observed between our datasets can be (a) reliably attributed to such long-term diachronic processes, which are usually observed over centuries, and (b) explained by such processes alone. The limitations of the approach we have taken, with respect to applying such theories to short-term diachronic change, are discussed by Aijmer (2018: 90), who notes, in the case of intensifier *fucking*, that short-term temporal differences are liable to idiosyncratic sociolinguistic variation of the time (see also Tagliamonte, 2008: 362). There is also macrostructural variation at play, even within relatively loosely-defined registers like casual conversation, as recently revealed by McEnery et al., (2023). Furthermore, our focus on teenage speakers, motivated by the broadly observed weakening of ꜰᴜᴄᴋ in the perception of young people, does not allow us to take into consideration age grading, as "the high frequency of fucking is correlated with a particular phase of life and this tendency is repeated in successive generations" (Aijmer, 2018: 75).

There are also limitations in terms of the data we have analysed. As discussed in Section 3, COLT and the Spoken BNC2014 were not compiled under the same circumstances. Of most critical relevance is that, while all participants in Spoken BNC2014 conversations were aware that they were being recorded, the level of awareness among speakers in COLT is variable (Love, 2020). This may well lead to uncontrolled variation in the use of taboo language. The broader context informing our choice of data is that, relative to written corpora, there are relatively few large, spoken corpora, and we are only at the beginning of "a new era in short-term diachronic corpus-based studies of spoken English" that has "lagged behind" when compared to diachronic studies of written corpora (Aijmer, 2022: 11). As such, future diachronic research into swearing (among many other linguistic phenomena) will benefit from the compilation of more datasets, sampling more periods in time, and facilitating more reliable comparison.

Finally, while our application of McEnery and Xiao's (2004) swearing function scheme for ꜰᴜᴄᴋ proved a useful window through which to infer pragmatic variability in our data and, after adaptation, with a high degree of inter-rater reliability, we recommend a further adaptation to the scheme in future research. As discussed in Section 3, in operationalizing the scheme we resorted to defining a closed class of idiomatic structures to be coded as category I. This had the benefit of ensuring mutual exclusivity between the function categories, but in doing so we adopted a view that the idiom status of ꜰᴜᴄᴋ is less a matter of (pragmatic) function and more a matter of form. In future research, it is our view that a more appropriate approach would be to remove category I from the scheme entirely, and code all instances of ꜰᴜᴄᴋ for idiom status independently. This would take into account our observation that idiomatic ꜰᴜᴄᴋ can be used to perform a range of functions that are not inherently dependent on idiom status (for example, *fucking hell* could be classified as an idiomatic form that functions as a general expletive). Had we adopted this approach in the present study, rather than undermine our observation that idiomatic ꜰᴜᴄᴋ has increased in usage (at the expense of other 'functions'), it would have allowed for the consideration of additional, partially idiomatized structures, and a more nuanced understanding of the functional role(s) of idiomatic ꜰᴜᴄᴋ among those functions that remain.

Stapleton et al. (2022: 8) argue that "much everyday use of swearing is not aimed at aggression, impoliteness, or even the expression of negative feelings". Using the terminology of Andersson and Trudgill (1990: 64), the use of ꜰᴜᴄᴋ among teenagers is highly *social* and *stylistic* and unlikely to be *abusive*, which appears to reflect a weakened taboo value of ꜰᴜᴄᴋ, considering its frequent use in amicable conversations between teenagers, where FUCK and combinations play an important role generally in adding to the friendly atmosphere. If it is indeed the case that the taboo status of ꜰᴜᴄᴋ is weakening among younger speakers, as indicated by perceptual research studies, then this study offers an empirical perspective on what the weakening of ꜰᴜᴄᴋ may 'look like' in real-time. While acknowledging that "a purely inductive inquiry of […] pragmatic variation solely dependent on a corpus renders the findings limited for broader generalization." (Pan, 2022: 410), we hope our findings may contribute to existing knowledge of and approaches to the forms, frequencies, functions and perceptions of ꜰᴜᴄᴋ. While the long-term

theoretical processes at play require further investigation, there can be little doubt that FUCK is a common and pragmatically valuable swear word among teenagers in British English.

Credit author statement

Robbie Love: Methodology, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing — Original Draft, Writing — Review & Editing, Visualization, Project administration. **Anna-Brita Stenström:** Conceptualization, Methodology, Resources, Writing — Original Draft, Supervision.

Declaration of competing interest

None.

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pragma.2023.08.008.

References

Aijmer, Karin, 2018. ''That's well bad': some new intensifiers in spoken British English'. In: Brezina, Vaclav, Love, Robbie, Aijmer, Karin (Eds.), Corpus Approaches to Contemporary British Speech Sociolinguistic Studies of the Spoken BNC2014. Routledge, New York, pp. 60—75.

Aijmer, Karin, 2022. 'Did you just basically steal everything?' e a study of discourse-pragmatic variation and change. J. Pragmat. 200, 8e23.

Andersson, Lars-Gunnar, Peter, Trudgill, 1990. *Bad Language*. Blackwell.

Beers Fägersten, Kristy, 2007. A sociolinguistic analysis of swearword offensiveness. Saarl. Work. Pap. Linguist. 1, 14—37.

Beers Fägersten, Kristy, 2017. Fuck cancer. Fucking Åmål, Aldrig fucka upp: the standardisation of FUCK in Swedish media. In: Advances in Swearing Research, vols. 87—106. Benjamins, Amsterdam.

BNC Consortium, 2007. The British National Corpus, XML Edition. Oxford Text Archive. http://hdl.handle.net/20.500.12024/2554.

Brezina, Vaclav, 2018. Statistics in Corpus Linguistics: A Practical Guide. Cambridge University Press, Cambridge.

Brinton, Laurel J., 1996. Pragmatic Markers in English: Grammaticalization and Discourse Functions. De Gruyter Mouton.

Brinton, Laurel, J., Traugott, Elizabeth Closs., 2003. *Lexicalization and Language Change*. Cambridge University Press.

Brown, Penelope, Levinson, Stephen C., 1987. Politeness: Some Universals in Language Usage. Cambridge University Press, Cambridge.

Bybee, Joan L., Pagliuca, William, 1985. Cross linguistic comparison and the development of grammatical meaning. In: Fisiak, Jacek (Ed.), Historical Semantics, Historical Word Formation. Trends in Linguistics, Studies and Monographs, vol. 29. Mouton de Gruyter, Berlin, pp. 59—83.

Carter, Ronald, McCarthy, Michael, 2006. Cambridge Grammar of English: A Comprehensive Guide to Spoken and Written English Grammar and Usage. Cambridge University Press, Cambridge.

Clancy, Brian, 2022. Building a corpus to represent a variety of a language. In: O'Keeffe, A., McCarthy, M. (Eds.), The Routledge Handbook of Corpus Linguistics, second ed. Routledge, New York, pp. 62—74.

Cohen, Jacob, 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20 (1), 37—46.

Crowdy, Steve, 1993. Spoken corpus design. Lit. Ling. Comput. 8 (4), 259—265. https://doi.org/10.1093/llc/8.4.259.

Culpeper, Jonathan, 2011. Impoliteness: Using Language to Cause Offence. Cambridge University Press, Cambridge.

Daly, Nicola, Holmes, Janet, Newton, Jonathan, Maria Stubbe, 2004. Expletives as solidarity signals in FTAs on the factory floor. J. Pragmat. 26, 945—964. https://doi.org/10.1016/j.pragma.2003.12.004.

Drange, Eli-Marie Danbolt, Hasund, Ingrid Kristine, Stenström, Anna-Brita, 2014. "Your mum!"Teenagers' swearing by mother in English, Spanish and Norwegian. Int. J. Corpus Linguist. 19 (1), 29—59. https://doi.org/10.1075/ijcl.19.1.02dra.

Drummond, Rob, 2020. Teenage swearing in the UK. Engl. World-Wide 41 (1), 59—88. https://doi.org/10.1075/eww.00040.dru [pre-print].

D'Arcy, Alexandra, 2015. Stability, stasis and change: the longue durée of intensifica- tion. Diachronica 32 (4), 449—493. https://doi.org/10.1075/dia.32.4.01dar.

Hasund, Ingrid Kristine, Drange, Eli-Marie Danbolt, Stenström, Anna-Brita, 2014. The pragmatic functions of swearing by mother in English, Spanish and Norwegian teenage talk. In: Rathje, Marianne (Ed.), *Swearing in the Nordic Countries*. The Language Board's Conference Series, pp. 11—35.

He, Yi, 2017. Really good or real good? A corpus-based approach to the grammaticalization and delexicalization of real. Thought 2, 1—20. https://so06.tci-thaijo.org/index.php/thoughts/issue/view/9400.

Heine, Bernd, Claudi, Ulrike, Hünnemeyer, Friederike, 1991. Grammaticalization: A Conceptual Framework. The University of Chicago Press, Chicago.

Herda, Damian, 2017. On metonymically motivated delexicalization of quantifying nouns in English and polish: a corpus investigation. Stud. Pol. Ling. 12 (4), 199—219. https://doi.org/10.4467/23005920SPL.17.010.8242.

Hopper, Paul K., Closs Traugott, Elizabeth, 2003. Grammaticalization, second ed. Cambridge University Press, Cambridge.

Ipsos MORI, 2021. *Public attitudes towards offensive language on TV and radio: summary report*. UK Office of Communications (Ofcom). https://www.ofcom.org.uk/__data/assets/pdf_file/0021/225336/offensive-language-summary-report.pdf.

Israel, Glenn D., 1996. Determining Sample Size. University of Florida. IFAS Extension, PEOD—6. https://www.psychosphere.com/Determining%20sample%20size%20by%20Glen%20Israel.pdf. (Accessed 12 April 2023).

Ito, Rika, Tagliamonte, Sali, 2003. Well weird, right dodgy, very strange, really cool: layering and recycling in English intensifiers. Lang. Soc. 32 (2), 257—279. https://doi.org/10.1017/S0047404503322055.

Jaffe, Alexandra, 2017. FUCK in French: evidence of "other-language" swearing in France and Québec. In: Fägrsten, Beers, Stapleton (Eds.), Advances in Swearing Research. Benjamins, Amsterdam, pp. 87—106.

Jay, Timothy, 1992. Cursing in America : a psycholinguistic study of dirty language in the courts, in the movies. In: The Schoolyards and on the Streets. John Benjamins, Philadelphia.

Jay, Timothy, 2009a. The utility and ubiquity of taboo words. Perspect. Psychol. Sci. 4, 152—161.

Jay, Timothy, 2009b. Do offensive words harm people? Psychol. Publ. Pol. Law 15 (No 2), 81—101.

Jay, Timothy, Janschewitz, Kristin, 2008. Pragmatics of swearing. J. Politeness Res. 4, 267—288.

Jay, Timothy, Janschewitz, Kristin, 2009. The utility and ubiquity of taboo words. Perspect. Psychol. Sci. 4, 152—161.

Kecskés, Istvan, 2000. A cognitive-pragmatic approach to situation-bound utterances. J. Pragmat. 32 (5), 605—625. https://doi.org/10.1016/S0378-2166(99)00063-6.

Kilgarriff, Adam, Baisa, Vít, Bušta, Jan, Jakubíček, Miloš, Kovář, Vojtěch, Michelfeit, Jan, Rychlý, Pavel, Suchomel, Vít, 2014. The Sketch engine: ten years on. Lexicography 1, 7—36.

Koester, Almut, 2022. Building small specialised corpora. In: O'Keeffe, A., McCarthy, M. (Eds.), The Routledge Handbook of Corpus Linguistics, second ed. Routledge, New York, pp. 48—61.

Lafreniere, Katherine C., Moore, Sarah G., Fisher, Robert J., 2022. The power of profanity: the meaning and impact of swear words in word of mouth. J. Market. Res. 59 (5), 908—925. https://doi.org/10.1177/00222437221078606.

Landis, J. Richard, Koch, Gary G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159—174.

Larsson, Tove, Paquot, Magali, Plonsky, Luke, 2020. Inter-rater reliability in Learner Corpus Research: insights from a collaborative study on adverb placement. Int. J. Learn. Corpus Res. 6 (2), 237—251. https://doi.org/10.1075/ijlcr.20001.lar.

Ljung, Magnus, 2007. Pragmatics in British Swearing. Paper presented at 28th International Computer Archive of Modern and Medieval English conference.

Ljung, Magnus, 2011. Swearing. A Cross-Cultural Linguistic Study. Palgrave Macmillan, Houndsmill.

Love, Robbie. 2020. *Overcoming Challenges in Corpus Construction: The Spoken British National Corpus 2014*. Routledge.

Love, Robbie. 2021. Swearing in informal spoken English: 1990s — 2010s. Text Talk 41 (5-6), 739-762.

Love, Robbie, Claire Dembry, Andrew Hardie, Vaclav Brezina, & Tony McEnery. 2017. The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. International Journal of Corpus Linguistics 22 (3), 319-344.

Lorenz, Gunter, 2002. *Really worthwhile* or *not really significant?* A corpus-based approach to the delexicalization and grammaticalization of intensifiers in Modern English. In: Wischer, Ilse, Gabriele Diewald (Eds.), New Reflections on Grammaticalization. John Benjamins, Amsterdam, pp. 157—176.

Marsden, Elizabeth, 2009. *What the* FUCK? *An Analysis of Swearing in Casual Conversation*. BA dissertation.

Mateo, José, Ramos Yus, Francisco, 2013. Towards a cross-cultural pragmatic taxonomy of insults. J. Lang. Aggress. Confl. 1 (1), 87—114. https://doi.org/10.1075/jlac.1.1.05mat.

McEnery, Tony, 2006. Swearing in English: Bad Language Purity and Power from 1596 to the Present. Routledge.

McEnery, Tony, Brookes, Gavin, Hanks, Elizabeth, Gerigk, Kevin, Jesse, Egbert, 2023. Swearing, discourse and function in conversational British English. J. Pragmat. 213, 36—48. https://doi.org/10.1016/j.pragma.2023.05.017.

McEnery, Tony, Xiao, Richard, 2004. British English: the case of FUCK in the BNC. Lang. Lit. 13 (3), 325—368.

McEnery, Tony, Baker, Paul, Andrew, Hardie, 1999. Assessing claims about language use with corpus data — swearing and abuse. In: Kirk, J.M. (Ed.), Corpora Galore: Papers from ICAME 1998. Rodopi, Amsterdam, pp. 45—55.

McEnery, Tony, Baker, Paul, Andrew, Hardie, 2000. Swearing and abuse in modern British English. In: Lewandowska-Tomaszczyk, B., Melia, P.J. (Eds.), PALC '99: Practical Applications in Language Corpora. Peter Lang, Frankfurt am Main, pp. 37—48.

Millwood-Hargrave, Andrea., 2000. Delete expletives? Research undertaken jointly by the Advertising Standards Authority. British Broadcasting Corporation, Broadcasting Standards Commission and the Independent Television Commission.

Mohr, Melissa, 2013. Holy Sh*t A Brief History of Swearing. Oxford University Press, Oxford.

Moinester, Murray, Gottfried, Ruth, 2014. Sample size estimation for correlations with pre-specified confidence interval. Quant. Meth. Psych. 10 (2), 124—130.

Pan, Yun, 2022. Intensification for discursive evaluation: a corpus-pragmatic view. Text Talk 42 (3), 391—417. https://doi.org/10.1515/text-2020-0046.

Partington, Alan, 1993. Corpus evidence of language change: the case of the intensifier. In: Baker, Mona, Francis, Gill, Tognini-Bonelli, Elena (Eds.), Text and Technology: in Honour of John Sinclair. John Benjamins.

Stapleton, Karyn, 2010. Swearing. In: Locher, M.A., Graham, S.L. (Eds.), *Interpersonal Pragmatics* (Handbooks of Pragmatics 6). De Gruyter, pp. 289—306.

Stapleton, Karyn, Fägersten, Kristy Beers, Stephens, Richard, Loveday, Catherine, 2022. The power of swearing: what we know and what we don't. Lingua 277, 1—16.

Stenström, Anna-Brita, Gisle Andersen, Ingrid Kristine, Hasund, 2002. *Trends in Teenage Talk: Corpus Compilation, Analysis and Findings*. John Benjamins.

Stenström, Anna-Brita, 2006. Taboo words in Teenage Talk: London and Madrid Girls' Conversations Compared. Spanish in Context 31, 115—138.

Stenström, Anna-Brita, & Leiv Egil Breivik. 1993. The Bergen Corpus of Teenage Talk. ICAME Journal 17, 128. https://icame.info/icame_static/archives/No_17_ICAME_Journal_index.pdf

Stenström, Anna-Brita. 2017. Swearing in English and Spanish teenage talk. In: Kristy Beers Fägersten & Karyn Stapleton (Eds.), Advances in Swearing Research. John Benjamins, pp. 157-183.

Stenström, Anna-Brita. 2014. Teenage Talk. From General Characteristics to the Use of Pragmatic Markers in a Contrastive Perspective. Palgrave Macmillan.

Stubbs, Michael, 2002. Two quantitative methods of studying phraseology in English. Int. J. Corpus Linguist. 7 (2), 215. https://doi.org/10.1075/ijcl.7.2.04stu, 214.

Tagliamonte, Sali A., 2008. So different and pretty cool! Recycling intensifiers in Toronto. Canada. Engl. Lang. Ling. 12 (2), 361—394. https://doi.org/10.1017/S1360674308002669.

Thelwall, Mike, 2008. Fk yea I swear: cursing and gender in MySpace. Corpora 3 (1), 83—107.

Tognini-Bonelli, Elena, 2000. Corpus classroom currency. Darbai ir Dienos 24, 205—243.

Torgersen, Eivind, Gabrielatos, Costas, Hoffmann, Sebastian, Fox, Susan, 2011. A corpus-based study of pragmatic markers in London English. Corpus Linguist. Linguistic Theory 7 (1), 93—118. https://doi.org/10.1515/cllt.2011.005.

Traugott, Elizabeth Closs, 1988. Pragmatic strengthening and grammaticalization. In: Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society (1988), pp. 406—416.

Traugott, Elizabeth Closs, 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. Language 65 (1), 31—55. https://doi.org/10.2307/414841.

Traugott, Elizabeth Closs, König, Ekkehard, 1991. The semantics-pragmatics of grammaticalization revisited. In: Heine, Bernd (Ed.), Elizabeth Closs Traugott, Approaches to Grammaticalization. Volume 1: Focus on Theoretical and Methodological Issues. John Benjamins, Amsterdam, pp. 198—218.

Van Olmen, Daniel, Marta Andersson, Jonathan Culpeper, 2023. Inherent linguistic impoliteness: The case of insultive YOU+NP in Dutch, English and Polish. J. Pragmat. 215, 22—40. https://doi.org/10.1016/j.pragma.2023.06.013.

Wajnryb, Ruth, 2004. Language Most Foul. Allen & Unwin.

Zenner, Eline, Ruette, Tom, Devriendt, Emma, 2017. In: Beers, Fägersten, Stapleton (Eds.), The borrowability of English swearwords: an exploration of Belgian Dutch and Netherlandic Dutch tweets, pp. 107—136.

Robbie Love is Lecturer in English Language at Aston University, UK. His research interests include the use of corpus methods to study contemporary spoken British English. He is the author of *Overcoming Challenges in Corpus Construction* and co-editor of *Corpus Approaches to Contemporary British Speech.*

Anna-Brita Stenström is Professor Emerita of Foreign Languages at the University of Bergen, Norway. She has initiated three online corpora of teenage talk in English, Nordic and Spanish. Her publications include *An Introduction to Spoken Interaction* and *Trends in Teenage Talk* (co-authored).