Fk yea I swear: Cursing and gender in a corpus of MySpace pages[1]

Mike Thelwall
School of Computing and Information Technology, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1SB, UK. E-mail: m.thelwall@wlv.ac.uk
Tel: +44 1902 321470 Fax: +44 1902 321478

Youth-orientated social network sites like MySpace are important venues for socialising and identity expression. Analysing such sites can therefore give a timely window into otherwise hidden aspects of contemporary culture. In this article a corpus of MySpace member home pages is used to analyse swearing in the U.S. and U.K. The results indicate that most MySpaces of 16 year olds and about 15% of middle-aged MySpaces contain strong swearing, for both males and females. There was no significant gender difference in the U.K. for strong swearing, especially for younger users (16-19). This is perhaps the first large-scale evidence of gender equality in strong swearing frequency in any informal English language use context. In contrast, U.S. male MySpaces contain significantly more strong swearing than those of females. The U.K. female assimilation of traditionally male swearing in the informal context of MySpace is suggestive of deeper changes in gender roles in society, possibly related to the recent rise in 'ladette culture'.

Keywords: Swearing, gender, MySpace, social network sites.

1. Introduction

MySpace is a highly popular youth oriented social network site, which apparently became the most visited web site in 2007 for U.S. web users (Prescott 2007). It has been described as replacing the (shopping) mall as the place where modern U.S. teens hang out (boyd in press). MySpace also has a wide user base in the U.K. (Prescott 2007) although other social network sites like Facebook and Hi5 are preferred by some population segments and in some countries. As a virtual place where many teens spend much time and express their identity with relative freedom, it is both an important aspect of youth culture and a place that should reflect offline modern youth attitudes and behaviour.

In this article the focus is on swearing. Swearing is interesting in itself as an aspect of language that is typically not taught or received from authority figures (parents, schools, politicians). It is also strongly tied to gender roles and expectations in society, and the words themselves are typically related to taboo issues. Hence, an analysis of swearing can be expected to provide a window into deeper social issues, particularly those relating to gender.

There has been widespread research into swearing within disciplines such as linguistics (McEnery 2005), neurolinguistics (van Lancker and Cummings 1999), psycholinguistics/developmental psychology (Jay 1992, 2000), sexuality (Sigel 2000), education (Dewaele 2004), history (Smith 1998), sociology (Stokoe and Edwards 2007), social psychology (Green 2003), women's studies (Bell and Reverby 2005), and nursing (Schapiro 2002). Nevertheless, it is difficult to gather large-scale natural language swearing data because written texts are normally produced in language registers that exclude swearing (McEnery and Xiao 2004). Hence considerable effort is needed to collect sufficient data to investigate factors such as social class, age and gender. For example, Jay's (1992:139) study of natural swearing was apparently the first extensive field study, using six travelling students to gather data on conversations involving a wide variety of people. Moreover, most psychological and sociological swearing research has used relatively small numbers of people from a restricted social group (often students). For example a recent study introduced a new

[1] This is an extended version of the following paper, incorporating extra background and literature review, plus the ambiguous words in Table 4 are excluded from the published version: Thelwall, M. (2008). Fk yea I swear: Cursing and gender in a corpus of MySpace pages, *Corpora*, 3(1), 83-107.

sampling method with an intermittently-activated electronic tape recorder gathering 30 second bursts of sound from the daily lives of 52 undergraduates, including some swearing (Mehl and Pennebaker 2003). As a result of the difficulty in gaining large-scale data, much linguistic swearing research has been qualitative or has used standard general purpose corpora of spoken texts, such as the British National Corpus (BNC) collection of conversations from 5,300 speakers gathered by 73 males and 75 females from the years 1991-1993 (Burnard 1995; Rayson, Leech, and Hodges 1997). Such standard corpora, which enjoy long term use and value, may not reflect current language trends (see Jay 1992:115-116 for a more detailed critique).

The web has been noted as an important source for corpus linguistics (Meyer, Grabowski, Han, Mantzouranis and Moses 2003) and the contents of social network sites like MySpace may potentially revolutionise swearing research because they are a relatively easily accessible source of large scale current linguistic data relevant to swearing. This is because the language of social network sites seems to be very informal and is probably closer to speech than most written forms. Moreover, a significant proportion of the population of some countries now uses social network sites and so sampling these sites gives access to a broad cross-section of people, although probably biased towards younger people and those with regular access to the Internet. Of course, social network language has some unique aspects – such as unusual spellings and acronyms, probably partly inherited from other electronic forms like instant messaging – and differs from normal conversation, for example due to the rarity of disfluencies like *er* and *um*, and probably also due to different conversation topics and rhythms of discussion. Nevertheless, large scale analyses of social network sites can be used to research social network language, which is important for its own sake, and also to investigate factors thought to be significant in contemporary swearing and other language use, such as gender, age, and social class.

Previous research in the U.K. and U.S. suggests that gender and age (and social class in the U.K.) are important factors in the propensity to swear and the type of swearing used. In particular, men seem to swear more than women (Bailey and Timm 1976) and males swear especially in all-male groups (Coates 2003; Bayard and Krishnayya 2001). In fact, recent U.K. research suggests that men use strong swear words more frequently than women, although women use milder swearing more (McEnery 2005). Personal observation suggests that this is no longer true; young women in the U.K. seem to swear more than men, even with the strongest swear words. This is an important issue not only for researchers but also for parents: if they glimpse their child swearing online, should they be concerned or has this become normal? Parents may be particularly concerned if their daughter swears (e.g., Jay 1992:32). The objective of this article is to test the gender gap hypothesis with current data from MySpace, including age as a likely interacting factor. The data used consists of a corpus of MySpace home pages from the U.K. and U.S.: the site's two biggest national user groups.

2. Swearing in English

Swearing is the use of any word or phrase that is likely to cause offence when used in middle class polite conversation (adapted from McEnery 2005). The emphasis here is on the language used rather than the content and so the definition excludes discussions of taboo topics when the language itself would not be recognised as the cause of offence. Perhaps in contrast to popular perceptions, however, the offensive nature of a word is unrelated to its sound but is socially constructed and changes over time. Swearing originally meant taking a legal or religious oath. The subsequent trivialising of such oaths (e.g., by God this meal is good) allowed the term swearing to take on a second meaning, one of using offensive language. The closely related term *curse* was initially used to refer to a wish or demand for something bad to happen (i.e., imprecate) but is now synonymous with *swear*. Similarly, *blasphemy* has purely religious origins but now has one meaning synonymous with swearing – in England this blurring apparently occurred in the nineteenth century, perhaps around high profile blasphemy trials (Marsh 1998:204-215). Other similar terms include: *bad, foul, Billingsgate, vulgar* or *coarse* language; *obscenity; profanity; oath; expletive; naughty* or

rude words. In addition there are many associated informal sayings such as *effing and blinding*, and to *swear like a fishwife/lord/ sailor/trooper*.

Swearing, including very mild words like *poo*, seems to be something that is a natural part of child development and children's swear words and the way they use them seem to depend upon their cognitive and social development stages, for example with focuses on potty training and sex differences at different ages (Jay 1992:35). Initially at least, swearing seems to perform the positive function of replacing violent actions as a means of expressing anger (Goodenough, cited in Jay, 1992:21), and many argue that swearing thus performs a useful role and should even be encouraged in context (Andersson and Trudgill 1992). Punishment seems to have little effect in any case (Jay, King, and Duncan 2006).

Historically, swearing in the U.K. has undergone major changes in the types of swearing employed, the types of words used, the actual words used, and in the social conventions about the contexts in which swearing is acceptable. Two extremes of acceptability in behaviour are a 1604 example of an English king giving "a turd for" a bishop's argument at an important theological meeting and being applauded for good debating skills (McEnery 2005:62); and a Victorian (late nineteenth century) polite society when referring to trousers was considered indelicate, even in private, and the euphemism *inexpressibles* used instead (Marsh 1998:215-230). The former example illustrates a time – possibly the majority of human history – when swearing did not particularly stigmatise the swearer even if, by definition, it shocked. Victorian society, in contrast, frowned upon swearing and stigmatised swearers as lower class and uneducated – a state of affairs that was partly the culmination of middle class campaigns against swearing and linking education to language moderation (McEnery 2005). During the twentieth century, however, swearing in public became gradually more acceptable. For instance, the first use of *bloody* on the U.K. stage in modern times was in the play *Pygmalion* in 1914 but swearing on stage is now unrestricted.

The following sections describe a range of categorisations and factors that can be used to dissect swearing.

1. Referents The types of offensive words or phrases used have changed over time, for instance with a reduction in religious and an increase in sexual connotations. Swearing today tends to make reference to current and past taboo subjects including: religion; sex acts; sexuality; genitals and sexual attributes; excretion; race, ethnic group or nationality; political affiliation (e.g., *commie*); any other denigrated or oppressed group (e.g., disabled, unemployed, old, young); stupidity; undesirable behaviour (e.g., *bitch, cow*); disease (e.g., *pox*). See also Montagu (Montagu 1967:chapter 6) for subcategories of most of these groups. Xenophobic and political words are probably the fastest-changing, and during times of crisis could even include the nationality of the opponents, although slang words could also be developed for such swearing. It is not necessary that the speaker or listener understands the connotation (e.g., *drat*, *pillock*, *git*), only that it is recognised as being potentially offensive. Note also that swearing does not necessarily have to include a word that is unambiguously a swear word. For instance, the last two words in "You bloody cow!" are swearing in this context but also have non-swearing meanings.

2. Linguistic types McEnery (2005:32) has categorised the following 15 linguistically distinct forms of swearing, based upon an analysis of the spoken section of the BNC (slightly paraphrased below).

- Predicative negative adjective: *the film is shit*
- Adverbial booster: *Fucking marvellous*
- Cursing expletive: *Fuck you!*
- Destinational usage: *Fuck off!*
- Emphatic adverb/adjective: *He fucking did it*
- Figurative extension of literal meaning: *to fuck about*
- General expletive: *Oh fuck!*
- Idiomatic set phrase: *fuck all*
- Literal usage denoting taboo referent: *We fucked*

- Imagery based on literal meaning: *kick the shit out of*
- Premodifying intensifying negative adjective: *the fucking idiot*
- Pronominal form with undefined referent: *got shit to do*
- Personal: Personal insult referring to defined entity: *You fuck!*
- Reclaimed usage with no negative intent, e.g., *Niggaz* as used by Black rappers
- Religious oath used for emphasis: *by God*

Some of the above linguistic forms associate with particular social classes or age groups. For instance, religious oaths seem to be the preserve of the older generation in the U.K., and literal uses with a taboo referent seem to be an upper class form of swearing (McEnery 2005:48-51). Most importantly here, all except reclaimed these forms of swearing seem primarily conversational rather than written. For example, variants of *fuck* in the BNC occur about thirteen times more frequently in spoken rather than written communication (McEnery and Xiao 2004). Moreover, fuck variants are almost non-existent in the more formal "context-governed" part of the BNC, with *fuck* occurring 150 times more often, per word, in dialogs in comparison to monologs (McEnery and Xiao 2004). Within written forms swearing can also relate to conversation: within in a novel, swearing would probably occur mainly in dialog. Pornography (Sigel 2000) and toilet wall and other graffiti (Green 2003) are exceptions, however, and swearing is occasionally found in other contexts, such as newspapers and magazines – particularly men's lifestyle magazines (Benwell 2001).

3. Word formation Swear words can be used on their own or in portmanteau formations, such as *motherfucker, shitfaced* and *cocksucker*. Mid-word interjections, such as *abso-bloody-lutely* and *kanga-bloody-roo* (Hughes 1991:24) also occur.

4. Purpose A common use of swearing is probably as an instinctive source of emotional release in response to sudden pain or bad news, but as the linguistic types above illustrate, cursing performs many services, many of which are not emotional. In most non-instinctive contexts, a swear word could be replaced by a milder synonym, so the speaker could have a specific conscious or unconscious purpose for swearing. This reason for swearing could be medical, such as brain damage or Tourrette's (van Lancker and Cummings 1999), although continual swearing could be a person's normal pattern of speech. In other cases, the decision to swear might have the purpose of expressing identity (e.g., being cool), or group membership or displaying closeness in friendship (Coates 2003). Swearing can also be used simply to communicate about taboo subjects, particularly for children (Jay 1992). Swearing occurs in jokes or for humorous intent (Andersson and Trudgill 1992, chapter 3; Liladhar 2000) and many jokes probably depend upon the shock value of swear words for their humour. Swearing in pornography works differently, being used to create an erotic effect (Sigel 2000). Finally, swearing is sometimes part of an organised ritual, as in flyting contests, including Black American *playing the dozens* (Abrahams 1962).

5. Strength Swearing varies in force, for example as measured by the percentage of people that would take offence at a particular usage. The level of offence of any given word seems to decline with usage (with the possible exception of fuck: Jay 1992). For example, *bloody* in the U.K. was shocking in the early 20th century but is mild today. Perceived strength also seems to be instinctive rather than a reasoned decision based upon word semantics (e.g., Dewaele 2004). In line with Wittgenstein's theory of rough language (Blair 2006), it is probably impossible to reliably differentiate between mild swearing and slang (e.g., *fatty*, *beanpole*), particularly when the slang is used in an abusive context.

6. Spellings Written swear words normally have a recognised official spelling, for example as given in the Oxford English Dictionary, although there are national variations in spelling and pronunciation of similar words. For example, U.K. English *shit* and *fuck* could be *shite* and *feck* in Ireland and both countries probably use both spellings to some extent. Moreover, written swear words may be bleeped (e.g., *sh*t*, *d---n you*, or even ****), accidentally miss-spelt (because they are rarely seen in print), deliberately spelt in dialect (e.g., *fook*) or just spelt in a cool or short way (e.g., *f0k, fok, fk*). A common form of 'misspelling' is probably to split portmanteau words or to join non-portmanteau words.

7. Implicit words Swearing is implicit when it is invoked by clever language. Examples include the brand name *FCUK*, the T-shirt slogan *Buck Fuddy*, abbreviated language such as *effing hell*, substituted words like *sugar* for *shit*, Cockney rhyming slang such as *Richard = Richard the Third = turd*, and euphemistic sexual humour (Lloyd 2007).

2.1 Gendered swearing in the U.K.

As introduced above, much previous research has found clear relationships between swearing and gender. As part of Victorian values in the U.K., women were considered too delicate to hear "bad language" and any sensitive topic could only be alluded to very indirectly. More recently, a general analysis of words used most frequently by one gender in conversation using the BNC conversation data from 1991-3 found that the word most distinctive of male speech was *fucking*, with *fuck* in seventh place (Rayson, Leech, and Hodges 1997). In contrast, there were no swear words in the top 25 most associated with female speech (Rayson, Leech, and Hodges 1997). A later analysis that focussed on swearing within the same data gave more detailed information. Women tended to swear as much as men, but using relatively mild language (e.g., *god, bloody, pig, hell, bugger*) so that men used more strong swearing than women (McEnery 2005). In a modern ironic twist on Victorian sexist politeness, today the use of swear words referring to female anatomy parts, can be regarded as sexist, especially when used by men in front of women (Dooling 1996; Stapleton 2003). In contrast, there are initiatives, such as the play *The Vagina Monologues*, to reclaim swear words for women by using them in a process of empowerment (Bell and Reverby 2005; see also Mills 1991). Moreover, there is a trend for younger women to swear more than older women, including with strong swear words (McEnery 2005).

In situations where a person is sworn at (i.e., abused), there are again gender differences, with males tending to swear at other males and targeting only relatively mild words at females. Females seem to abuse both males and females but abuse females more, with gender differences in the choice of words. For instance, only women seem to be called a *cow* and men are mainly *bastards* (McEnery 2005). Although there seems to be no specific evidence of this, it seems likely that same gender conversations use more swearing than mixed gender conversations, for example in conversations about sex (Hey 1997:80-83).

Why have men tended to swear more than women, at least over the past few hundred years in the U.K.? The military is sometimes invoked as a reason (Montagu 1967:109). Military service can bring together large bodies of men, which may promote standardised language. In times of stress, such as tough training and war, swearing can help to vent emotions and so it would be unsurprising to hear soldiers curse frequently. A weaker version of the same argument perhaps holds for collective work, such as in factories (e.g., Smith 1998), where the community of practice concept could apply (Stapleton 2003). Hence it is possible that gender differences in swearing are related to some extent with the proportion of women in employment and in the armed services.

Swearing differences in the U.K. are not just restricted to gender, there are also class differences, with lower-middle and upper-working class (i.e., social class C1 and C2) tending to swear much less than unskilled manual and unemployed (DE) speakers. Upper-middle and middle class (AB) speakers also swear somewhat less than lower class (DE) speakers but typically in different ways (McEnery 2005).

2.2 Gendered swearing in the U.S.

U.S. society can appear to be intolerant of swearing, for example with the Federal Communications Commission completely excluding "seven dirty words" (Sapolsky and Kaye 2005) from broadcast television (*shit, piss, fuck, cunt, cocksucker, motherfucker, tits*), whereas in the U.K. all swearing has been allowed for decades at appropriate times. Some swearing does occur in U.S. prime time broadcast television, however, with mild words such as *hell* being common and others occurring such as *butt, ass, screw, suck, bitch, son-of-a-bitch* and *bastard* (Sapolsky and Kaye 2005). Practices are changing, however, with U.S. cable TV producing successful series containing swearing, such as South Park (Grimm 2003). The relatively prudish nature of U.S. TV does not imply a prudish attitude in the population,

however. The TV swearing embargo is driven by advertising, whereas U.S. movies have exploited their ability to use swearwords freely (Jay 1992:222-230), after dropping the restrictive Hayes Code in response to tough TV competition from the 1950s (Hughes 1991:198-199). Moreover, swearing seems to be common in the U.S. (Jay 1992:155).

Gendered patterns of swearing have probably been similar historically in the U.S. to the U.K. A questionnaire-based study of U.S. students in 1974 found that women reported using more weak and less strong swear words than men (Bailey and Timm 1976). A detailed study of swearing in children (using covert recording) showed that gender differences start at age 5, with boys and girls having partly different vocabularies and boys developing more frequent swearing than girls. Anger and frustration were the main causes of swearing (Jay 1992:68). All U.S. studies of gender difference in swearing seem to have found males swearing more and with stronger words, and both genders swearing more in same sex groups (Jay 1992:169; 2000; see also Jay 1980a). Nevertheless, linguistic changes are occurring: for example whilst urban women's spoken language was closer to standard English than men's in the 1960s (Trudgill 1972), this seems to have recently reversed, at least for working class girls (Eckert 2003).

Most swearing research in the U.S. seems to have been conducted by experimental psychologists, but the results are not discussed here because they are not directly relevant (for a review and bibliography, see: Jay 1992, 2000).

The set of swear words used in the U.K. and U.S. are probably similar but with some differences. For example, the word *fanny* is a mild word for bottom in the U.S. whereas it is probably rarely used or understood in the U.K., although it is more common in Ireland where it can be very strong (Stapleton 2003). Similarly, the U.K. swear word *bollocks* seems little known in the U.S. There is one important example of a swear word, *motherfucker*, that was invented relatively recently (before 1956 by Black Americans) in the U.S. before spreading to the U.K., perhaps through film and music. Two strong swear words (*cocksucker* and *cockteaser*) seem to be common in the U.S. but rare in the U.K. (Jay 1992:162).

3. MySpace

MySpace, like all social network sites, is a web site that allows internet users to join and create their own home page (also called here 'a MySpace'). This page is likely to contain a photograph and personal details such as age, gender and personal interests. MySpace members are allowed to register with each other as friends, with members having pictures of friends on their home page. Below these pictures is a comments section used by registered friends to send messages. In fact, many members have conversations with friends by taking turns to write on each other's comments section. This is an unusually public way to have a conversation since anyone may read the comments.

In order to understand comment conversations it is important to keep in mind that they are also public *identity performances* (boyd and Heer 2006). Even for members with private profiles (including all under 16s) all comments received are visible to all their friends and are therefore relatively public. Teen MySpaces are identity performances because of this public nature. As part of this, the profile portrait is often cool or funny and comments are often interesting or amusing, with inventive use of language. A member may take care when writing comments in the knowledge that shared friends are likely to read them and judge them. Indeed, a member may believe that their skill at writing in MySpace gives them offline status amongst their peers (boyd 2006). In terms of swearing, this probably encourages the use of non-standard spellings and humorous contexts. Abusive and negative swearing may be relatively uncommon, since social network communication is between registered friends. This is a potential point of difference between the pragmatics of MySpace and spoken swearing.

Although MySpaces are normally public, it can be difficult to search for the MySpace of a known person because members often use pseudonyms. In this sense MySpace can be a kind of parallel identity and can be also insulated from discovery by authority figures (parents, teachers, employers) who may frown upon the language used. Nevertheless, parents sometimes seek to control the language and content of their offspring's MySpace, with a common reaction to this being to create a second, more personal MySpace that is hidden from

parents (boyd in press). Hence, whilst the language in some teen MySpaces may be censored, this is probably the exception rather than the norm.

Most academic research and media coverage of MySpace has focussed on teen users even though the median age of members may actually be about 21 (Thelwall, 2007). Moreover, there has been apparently no academic research or media coverage investigating swearing. This may be because academic research into the relatively new phenomenon of Web 2.0 social network sites has tended to focus on friendship as the key component (Fono and Raynes-Goldie 2005), to be quite holistic (e.g., Ellison, Steinfield, and Lampe 2006) or to specialise away from swearing (Golder, Wilkinson, and Huberman 2007). Nevertheless, as the results below illustrate, swearing is an important MySpace activity.

3.1 Swearing and identity in MySpace

Swearing in MySpaces can occur in several different places. Most prominently, the member name itself sometimes contains swear words. Second, from the perspective of the site owner, swear words can be inserted into the self-description and free text parts of the home page. For example many users have imported and answered a list of standard questions about themselves, including: "Do you swear?", which has many answers like: "Fk yea" or "Of course I bloody don't". The other main place in which swearing can be used is the MySpace blog, which has headings displayed on the home page but not the full blog entry. Hence the blog could contain swearing that is not visible from the home page. For example, one MySpace has only one swear word visible - *dickheads* – in the title of a blog posting containing several other swear words and describing the owner's car being stolen. Overall, since the MySpace is a kind of identity performance for the member they can take decisions with regard to the role of swearing in their identity. For example they could decide to keep their profile 'clean', to use a lot of swearing, or to ban certain words.

The friends' comments section is another place where swearing can occur on the user home/profile pages. These comments are not made by the member but they still have direct control over them in the sense that they choose their friends and can subsequently 'defriend' offenders if necessary. Perhaps most importantly, friends can be expected to make judgements about appropriate language for comments. Friends' comments are part of the host MySpace member's identity in this sense and also because of their location within the target member's MySpace. One previous paper has reported the extent of swearing of adolescents in MySpace (for members under 18 with public profiles), finding that 19.9% of profiles contained swearing in the owner-controlled text parts and 32.8% contained swearing in the friends' comments section.

4. Research Questions

The following research question is addressed in this article, although the research also has the secondary purpose of shedding light on wider gender differences in language use in society.

- Is there a gender difference in the proportion of male and female MySpace profiles allowing strong swearing in the U.K. and in the U.S.?

The choice of the U.K. and U.S. above is pragmatic, these being the two largest national users of MySpace.

5. Data Collection

The overall research design was to gather a corpus of MySpace home pages and to check each page for the occurrence of a set of known and common strong and moderate swear words and then to compare the occurrence of these words with the owner's gender, age and nationality. Although corpus approaches have been used to study other forms of computer-mediated communication (e.g., Ooi 2000, Payne 2005), there do not seem to be previous corpus analyses of social network sites. As discussed above, member home pages are taken as a whole as identity performances. In particular, friends' comments are exclusively associated with the recipients (i.e., the MySpace in which the comments appear). A small preliminary analysis suggested that about three quarters of strong swearing occurred in the comments

section, and that the gender of the swearing commenter did not strongly associate with the gender of the MySpace owner.

To gather a large sample of members, 40,000 who registered with MySpace about a year before data collection were chosen. Members joining at the same time were used in order to have a consistent sample. The large sample size was determined by the need to perform statistical tests on the U.K. subset, a relatively small proportion of all members. Members' joining times were determined by their ID number and their home pages were downloaded via the URL http://profile.myspace.com/index.cfm?fuseaction=user.viewprofile&friendid= followed by their ID. The IDs ranged from 90306349 to 90346348. The program SocSciBot (http://socscibot.wlv.ac.uk) was used to download the profile pages (without embedded images to minimise network and resource usage). The first 10,000 profiles are reused from a previous paper (Thelwall submitted) and were downloaded over four hours on July 17, 2007. The second batch of 30,000 was downloaded at a rate of about one every 10 seconds from August 3 to 7, 2007 to minimise impact upon the MySpace servers.

Each downloaded MySpace page was automatically checked and the following pages were discarded:

- Ex-members
- Pages with download or other general errors
- Private pages
- Members who had not declared their location as either the U.K. or U.S.
- Members with 0 or 1 friends, which are unlikely to be active users (Thelwall submitted)
- Film makers, comedians and musicians – these spaces are typically for the promotion of artists rather than being personal spaces

For the U.K. this gave a final data set of 767 MySpaces (281 female, 486 male) and for the U.S. 8,609 (3,950 female, 4,659 male). The discrepancy is due to MySpace being more popular in the U.S. and the U.S. being a larger country. Each page was parsed in full to extract a complete list of words used (after depluralisation) and a central vocabulary was created of all words used in all profiles. Note that the typical approach in previous swearing research was to start with a small sample of individuals and to record their swearing over a set period of time, performing statistical tests on the combined set of swear words (McEnery 2005; Rayson, Leech, and Hodges 1997) or to ask a small set of individuals about their swearing practices (Stapleton 2003). The former is not ideal because individual heavy swearers can have an undue influence on the results, giving spurious positive statistical tests. In contrast, we have sufficient numbers to count how many *people* are associated with a given group of swear words and perform statistical tests on a per-person rather than per-word basis. This makes the outcomes more robust.

Note that an unknown proportion of MySpace member profiles are misleading, with the member giving deliberately incorrect information. Probably the most common source of incorrect data is age, although some users also report an incorrect gender or impersonate a celebrity. The only available relevant statistic so far seems to be that about 8.3% of members with public profiles with ages under 17 contain evidence that the members are younger than reported (Hinduja and Patchin in press). A previous study has also suggested that the age profiles seemed to be broadly correct in the sense of giving reasonable overall picture of users (Thelwall submitted). Hence it seems that age fraud will affect a small percentage of the data, and probably under 10%. Anecdotal evidence suggests that gender switching, in particular, is most common with older males and hence should not significantly affect the MySpace data.

6. Swear word selection

The level of offence of an individual swear word depends upon the context in which used. According to the British Board of film Censors (BBFC), "it is impossible to set out comprehensive lists of acceptable words or expressions which will satisfy all sections of the public" (http://www.bbfc.co.uk/policy/policy-mainissues.php, accessed August 30, 2007). Nevertheless, there are certain words that are recognised as being offensive to a significant

number of people in almost irrespective of how they are used. In the current research, in order to assess the extent to which swearing is present in MySpaces, a simplistic approach was taken: to construct a list of swear words and then to count MySpaces containing them. This is simplistic because some swear words, particularly the weaker ones, are only thought of as swearing when used in the context of abuse. Other swear words are unambiguous, for example vulgar versions of conventional words. A different related issue is the use of reclaimed words. For example, a gay man using the word *queer* to describe himself gives the word a positive non-offensive context, whereas it could be highly threatening from a stranger. Finally, offensive words are regularly used in contexts where they are terms of endearment, for example friends calling each other *motherfucker* (Jay 1992:177) and hence, whilst used because they are offensive, are not intended to offend.

A list of swear words used was compiled from a combination of sources. First, the official British Broadcasting Corporation guidelines were used, as quoted in McEnery (2005). This gives the strength of offence of words on a five point scale from very mild to very strong. For each of these words common word variants were added, for example through common suffixes like –ed, -er, and –ing. In addition, portmanteau words were sought by a full-text search for word stems using the vocabulary (Thelwall and Price 2006), Next, the vocabulary was browsed to seek spelling variants for known swear words, whether spelling mistakes or deliberately unorthodox spellings. Finally a small number of additional known swear words were added. Portmanteau words were so frequent that not all could be added and so a word frequency restriction was used, excluding words occurring less often than once per 1,000 MySpaces.

The words in the list were classified by strength by extrapolating from the BBC classifications (Table 6). For the main analysis (tables 1-3) each ambiguous word was checked for the relative frequency of offensive and non-offensive uses, rejecting those without at least 75% offensive uses (without making judgements about whether the intention was to offend).

Finally, for each U.K. and U.S. MySpace in the corpus a count was made of the frequency of occurrence of the three strengths of swear words, using the Cyclist software associated with SocSciBot. This, together with the gender and reported age of each MySpace owner, is the raw data analysed.

Since there are variations in the usage and strength of swear words between English-speaking nations, a single set of swear words may be inappropriate for international studies. For the U.S., the list of words most equivalent to the BBC's is probably the seven dirty words discussed above that are excluded from broadcast TV. This is an authoritative list of reasonably strong swear words, although it seems to be flawed. For example racial swearing is not present in this list, *shit* and *piss* seem to be relatively mild U.S. swear words and *tits* is probably a medium-strength word (Jay 1992:148-151). A secondary analysis of the data was conducted for these words, for comparison purposes. In each case variants of the words were included, such as different spellings and portmanteau terms. The word *cocksucker* was not included because it fell below the minimum usage threshold, occurring only three times in the 40,000 profiles. Possibly it occurred more often as two separate words.

7. Results

7.1 Overall swearing

7.1.1 The BBC word list

Table 1 reports the main results: the percentage of profiles containing at least one word in each of the BBC-derived categories of at least moderate strength. Weaker swear words were not analysed here: these are problematic because many are ambiguous, and it seems impossible to get a large enough list to give reasonable coverage. Many were included in the identification stage (see Table 6) mainly as part of a process of identifying the moderate and strong swear words. In the U.K. male MySpaces contained significantly more moderate (Chi-Square, $p=0.001$) and very strong (Chi-Square, $p=0.014$) language but not significantly more

strong (Chi-Square, p=0.316) language. In contrast, U.S. male MySpaces contained significantly more strong (Chi-Square, p=0.000) language but not significantly more moderate (Chi-Square, p=0.127) and very strong (Chi-Square, p=0.577) language.

Note that "very strong" essentially equates to *cunt* or variations of *motherfucker* and "strong" equates to variations of *fuck* (see Table 6). This categorisation is also appropriate for the U.S., although the omitted low frequency words *cocksucker* and *cockteaser* could also be categorised as either strong or very strong in the U.S. (Jay 1992:162 - ratings from 1978).

Table 1. Percentage of profiles containing different strengths of swear words (BBC classification – see Table 6 for words in each category).

	Moderate	*Strong*	*Very strong*	*Sample size*
US Males	8%	*32%*	1%	4659
US Females	8%	*25%*	1%	3950
UK Males	*25%*	30%	*5%*	486
UK Females	*14%*	25%	*2%*	281

The *ratio* of swear words to total words in MySpace profiles seems to be relatively small (see Figure 1): combining the strong and very strong categories for the U.K., the maximum percentage of swear words was 5%, with the mean for swearers being 0.2%. Analysing just U.K. swearers, male MySpaces (0.23%) had a slightly higher proportion of swear words than female MySpaces (0.15%) but the difference between males and females was not statistically significant (Mann-Whitney U test). The U.S. was similar, with a maximum percentage of 11%, and a mean of 0.3%. Also analysing just U.S. swearers, male MySpaces (0.3%) had a significantly higher proportion of swear words than female MySpaces (0.2%) (Mann-Whitney U test, p=0.000). Other studies reporting swearing rates in informal spoken English have found a similar figure: about 0.5% (Jay 1980b, Mehl and Pennebaker 2003).

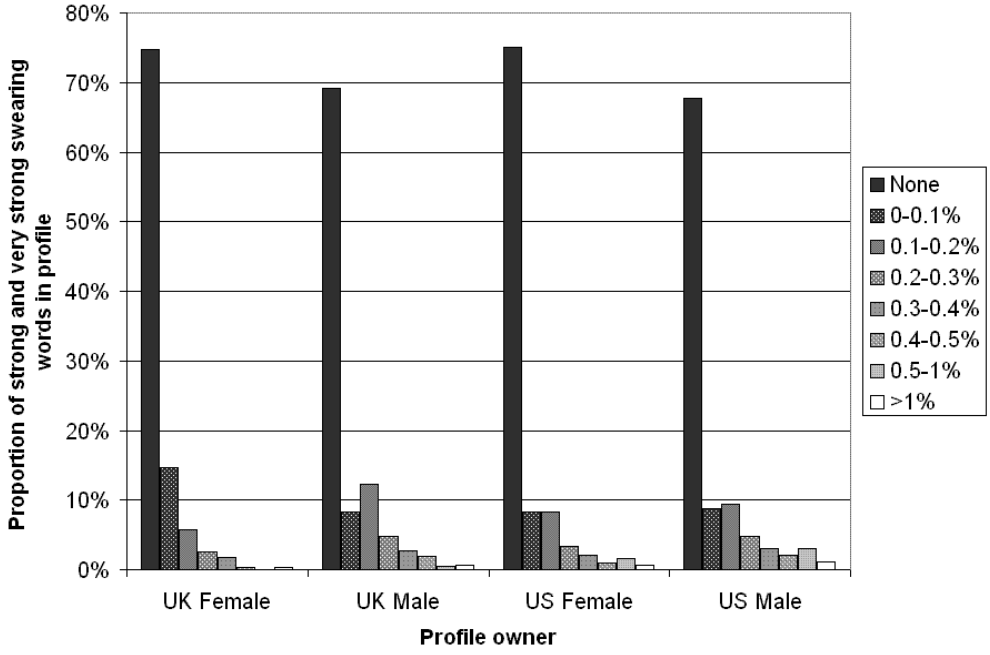

Fig. 1. The proportion of unambiguous strong and very strong swear words within profiles (all bold words in Table 6).

The analysis was repeated for just the profiles with owners reporting an age of 16-19, and the results shown in Table 2. Note that this reported age is not necessarily correct, although it seems likely to be approximately correct in the majority of cases because the overall profile of MySpace reported ages seems reasonable (Thelwall submitted). For U.K.

under-20s, male MySpaces contained significantly more moderate (Chi-Square, p=0.009) language but not significantly more strong (Chi-Square, p=0.393) or very strong (Chi-Square, p=0.167) language. For U.S. under-20s, male MySpaces contained significantly more strong (Chi-Square, p=0.000) language but not significantly more moderate (Chi-Square, p=0.199) and very strong (Chi-Square, p=0.700) language.

In all cases the younger users' MySpaces contained more swearing than average. Female use of swear words was greater than male use for younger users in two cases: moderate language in the U.S. and strong language in the U.K. – although this difference was not statistically significant in either case. Nevertheless, the findings for younger U.K. users strengthens the evidence that in the U.K., strong language is no longer dominated by males.

Table 2. Percentage of profiles containing different strengths of swear words – younger members (BBC classification).

	Moderate	*Strong*	*Very strong*	*Sample size*
US Males 16-19	10%	*47%*	2%	1530
US Females 16-19	11%	*38%*	2%	1287
UK Males 16-19	*33%*	33%	*8%*	171
UK Females 16-19	*18%*	38%	*3%*	130

7.1.2 The seven "dirty" words

The analysis for all age groups was repeated for the U.S. "seven dirty words" list, with a significant gender difference in the U.S. (males: 42%, females: 38%, p=0.000). This confirms the male bias for strong swearing in U.S. MySpaces.

7.2 Common words

Tables 3 and 4 report the most male-oriented and female-oriented swear words analysed, in terms of the difference in percentage of MySpaces containing these words. These words were not context-checked to exclude non-swearing uses because of the numbers involved (U.S.: 47,657; U.K.: 3,139). A continuity-corrected chi-square value is given, as well as an associated *p* value. Note, however, that for the calculations so the reported *p* values are for guidance only: they are underestimates because the tests summarise only the significant findings from 100 tests (i.e., no Bonferroni corrections are used). For example, it is likely that at least one of the words in the tables is incorrectly attributed a gender association in MySpace.

The U.K. list includes no unambiguous swear words with a statistically significant association with females (*hell* and *dirty* are ambiguous), although the female preference for *cow* is expected (McEnery 2005:39). The data confirms previous U.K. findings that two strong swear words, *fuck* and *cunt* are predominantly male words (but not for younger users 16-19), although *ass* is unexpected (McEnery 2005:35). Note that these results partly contradict the main findings for strong swear words: although *fuck* has a gender bias overall, once the variations of *fuck* are included (i.e., the strong swearing category of Table 1) the gender association is no longer significant.

Table 3. Percentage of profiles containing specific words, broken down by profile owner gender (U.K. MySpaces, words with significant gender differences only).

Word	*Female*	*Male*	*Chi-square*	*p*
cow	**5%**	2%	4.790	0.026
ass	7%	**13%**	4.665	0.031
cunt	1%	**5%**	5.791	0.016
fuck	9%	**15%**	5.795	0.016

For the U.S. words (Table 4), it is striking that the two unambiguous swear words (*whore*, *slut*) that are significantly more female than male are both female-oriented abusive words for promiscuity (Jay 1992:180), in contrast to the milder U.K. *cow*. The male swear

words cover many types, including sexual, racial, scatological, and religious (possibly mainly in non-swearing contexts). There are several variations of *fuck*, although surprisingly none are top of the list. The reclaimed spelling *nigga* is an interesting inclusion. Examination of a sample of MySpaces indicates that it is predominantly used by African-Americans, often very prominently, and its male-orientation in this context appears to be a significant cultural finding. Perhaps African-American women tend to avoid this term (although it is in 9.3% of U.S. female MySpaces) because of its racial origins, or perhaps its use by men is a positive statement of identity in a way that is not seen as necessary by women. Finally, the table confirms the rise of *gay* as a common U.S. swear word. Unlike *nigga*, it is predominantly used in abusive (non-reclaimed) contexts, although it is used in both ways. It is typically not used as an anti-homosexual insult but as a synonym for *stupid* or *bad* (there is a Facebook group called "The word "gay" is not a synonym for "stupid"", with 87,014 members on September 26, 2007), or to tease friends.

Table 4. Percentage of profiles containing specific words, broken down by profile owner gender (U.S. MySpaces, words with significant gender differences only).

Word	Female	Male	Chi-square	p
tart	**1.8%**	0.7%	23.255	0.000
god	**19.6%**	17.3%	6.923	0.009
slut	**2.0%**	1.4%	5.524	0.019
whore	**3.0%**	2.2%	5.151	0.023
dirty	**5.9%**	4.9%	4.673	0.031
butthead	0.3%	**0.6%**	3.932	0.047
screw	0.8%	**1.2%**	3.985	0.046
turd	0.2%	**0.4%**	3.992	0.046
jerk	0.7%	**1.2%**	4.882	0.027
retard	0.4%	**0.8%**	5.444	0.020
jew	0.3%	**0.7%**	5.971	0.015
hell	18.4%	**20.7%**	6.650	0.010
pussy	2.0%	**2.9%**	7.169	0.007
nigger	0.2%	**0.6%**	7.309	0.007
asshole	1.2%	**2.2%**	12.687	0.000
dick	2.5%	**3.8%**	12.796	0.000
queer	0.2%	**0.7%**	13.467	0.000
fucking	10.4%	**13.2%**	15.190	0.000
ass	22.9%	**26.6%**	15.696	0.000
shit	24.3%	**28.4%**	19.012	0.000
cock	0.5%	**1.5%**	20.858	0.000
fuck	15.4%	**19.7%**	26.951	0.000
fuckin	6.8%	**10.2%**	31.310	0.000
pimp	5.1%	**8.3%**	34.881	0.000
nigga	9.3%	**14.1%**	45.686	0.000
fucker	1.2%	**3.6%**	51.609	0.000
gay	4.5%	**10.0%**	90.923	0.000

7.3 Age factors

In general, swearing has been shown to decline with age and there is sufficient data for the U.S. MySpaces (up to age 40) to check that this holds in this online context. Figure 2 shows the percentage of MySpaces that contain at least one of the unambiguous strong or very strong words in Table 6, showing a similar pattern for males and females of declining swearing with age.

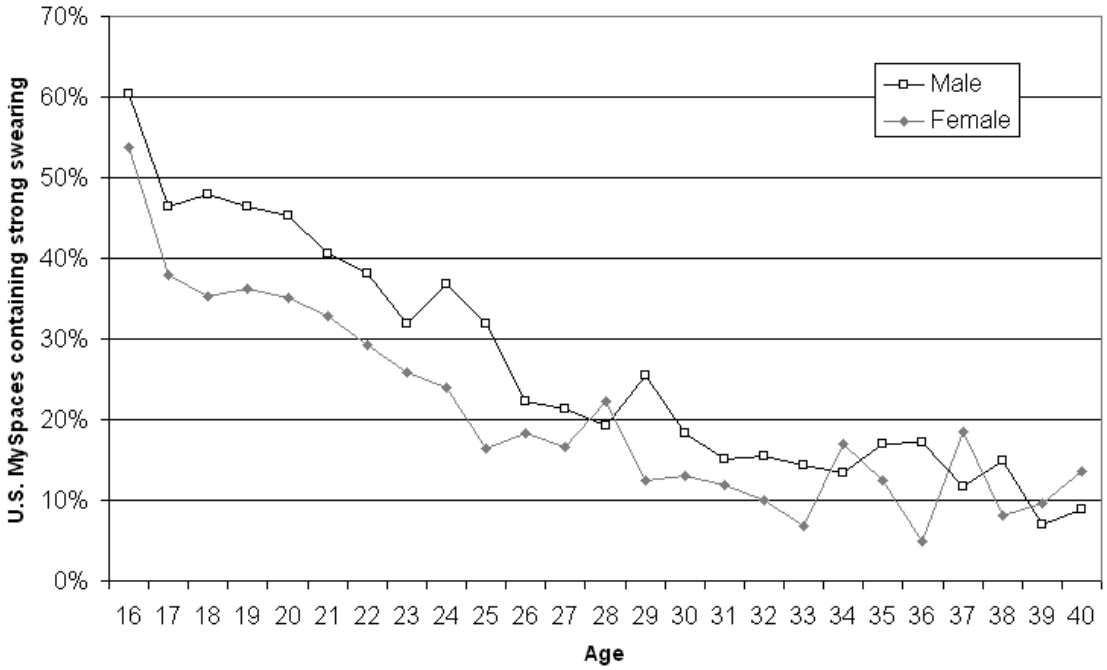

Fig. 2. A profile owner age and gender breakdown of all unambiguous strong and very strong swearing in U.S. MySpaces (all bold words in Table 6).

Figure 3 illustrates the typical age-related strong swearing pattern, using the example of *fuck*. In this case (and other examples not shown) the gender gap is small for the youngest users, but increases from age 17. It is clear however, that strong swearing in MySpace is normal for teenagers but rare for middle aged users, both male and female.

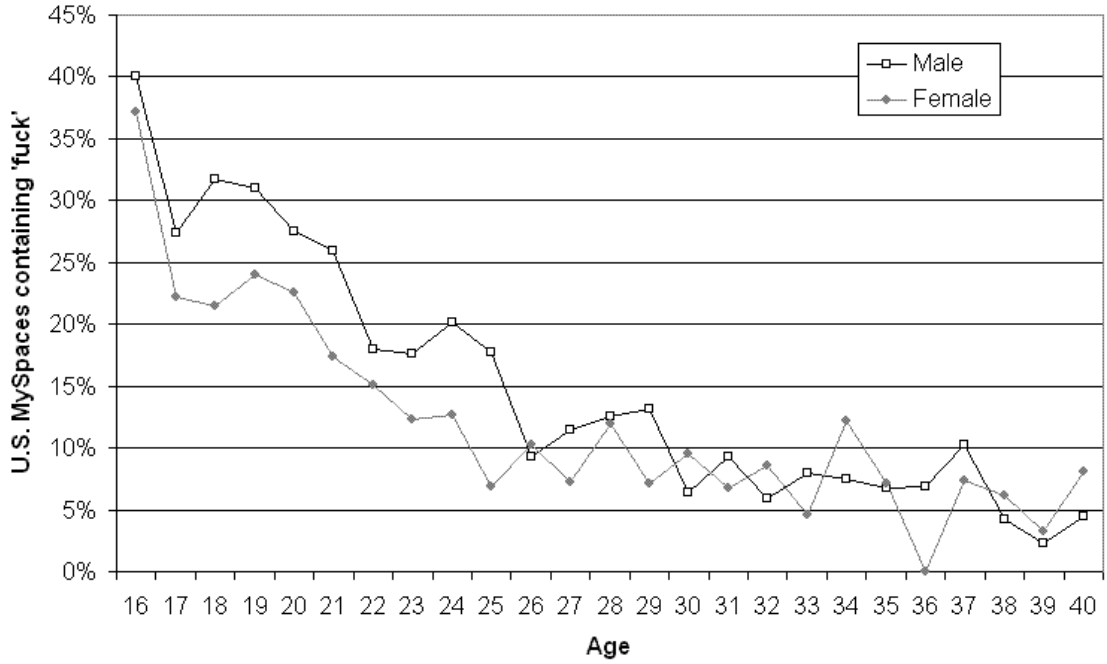

Fig. 3. A *fuck* profile owner age and gender breakdown in U.S. MySpaces.

7.4. British swearing types

In order to give some context to the results, Table 5 presents a classification of the swearing using McEnery's scheme. Following McEnery and Xiao (2004), some of the similar categories were merged. All identified swear words used in the main analysis (bold words in Table 6) in the U.K. MySpaces were classified. Note that McEnery's final two categories do

not apply because no religious words are included. The classifications were made by the author and a helper, with the author making the final decision in cases of disagreement. In some cases the category was an educated guess. In particular, the phrase *get fucked* in comments like "I'm going to get fucked at the weekend" was always interpreted as an idiomatic set phrase meaning *getting drunk* or *taking drugs* rather than *having sex* (i.e., literal usage denoting taboo referent).

One distinctive form of swearing was the implicit idiomatic phrase. For example, in: "aye so looks like coatbridge toon centre for me! its got a [shop name] and a [shop name]!! fuck you got? A [shop name]! Lolololololol" (lol is code for "laugh out loud"), "fuck you got?" was interpreted as short for "what the fuck have you got" and was coded as an idiomatic set phrase. Similarly "qu'est ce que fuck?" was interpreted as "what the fuck?" It was noticeable that most of the swearing was playful. Even the negative contexts for swearing were mostly directed at the author or ironic –many of them finished with a kiss "x" or some other way of making clear that the negative comment should not be taken seriously (e.g., "fuck off now and i mean that! x"). The personal insult category could easily be split into "ironic self-insult" and "playful pseudo insult or term of endearment", although such splits would be entirely based upon judgement rather than linguistic context. For example "twat face cunt hed!!!!!!!" is a playful friend comment although its linguistic form is of a pure insult. Interestingly, one comment focussed specifically on gender and swearing: "I can be quite sarcastic but funny with it! I am foulmouthed, but if you're gonna get any man to listen to you these days you have to talk like them,(mother fuckers!!)".

In comparison with a similar classification of *fuck* variants in the whole British National Corpus (McEnery and Xiao 2004), the main differences are: emphatic usage is lower in MySpace (58% in the BNC) but personal insults are much more common (2% in the BNC) and idiomatic set usage are also more common (12% in the BNC). Although the words covered are different in the two studies, *fucking* is common in the emphatic category and so it seems that swearing is relatively underemployed for emphasis in U.K. MySpaces, in comparison to other uses. The relative idiomatic/figurative emphasis in MySpace is also clear, including in the personal insults category, as discussed above. Finally, the inclusion of a significant amount of general expletives strongly emphasises that MySpace language can have features that are otherwise distinctive of spoken language.

The clearest gender difference in Table 5 is in male profiles containing 11% more idiomatic set phrases (or Figurative extensions of literal meanings) than females. This is partly due to more expressions meaning getting drunk, although these were present in both male and female cases.

Table 5. A classification of UK MySpace swearing types from all the UK MySpace users in the sample, classified by profile owner (427 swear words classified, 264 male and 163 female; all except 67 were friends' comments).

Type	% of all swearing M	F	Examples (slashes // separate comments)
Predicative negative adjective	0%	1%	..and your myspace page is fucked
Cursing expletive	1%	3%	yeah but. IVE JUST GOT BACK FROM A DEFTONES GIG! so fuck you jim bob! X // so id bollocks to this
Destinational usage	3%	1%	fuck....right....off.... // Chris you're slacking again !!! Get the fuck off myspace lol !! you good anyway ?
Emphatic adverb/adjective OR Adverbial booster OR Premodifying intensifying negative adjective	32%	38%	and we r guna go to town again n make a ryt fuckin nyt of it again lol // see look i'm fucking commenting u back // lol and stop fucking tickleing me!! // Thanks for the party last night it was fucking good and you are great hosts. // That 50's rock and roll weekender was fucking mint! // Fuckin my space, my arse // Fucking goths! // 1/2 d ppl cudnt even speak fuckin english! // yeah so me and sarah broke up and everythings fucking shit // sick of being hurt all the bastard time.
General expletive	6%	7%	fuck i didnt know this was u lol // lol.....nooooo waaaayyyyyy....fuck !!! // Ahh Fuck it, we're doomed anyway..
Idiomatic set phrase OR Figurative extension of literal meaning	28%	17%	think am gonna get him an album or summet fuck nows // got another copy of the reaction CD (will had fucked the last one lol) // qu'est ce que fuck? // what the fuck pubehead whos pete and why is this necicery mate // Heh long story.. cant be fucked to explain :D
Literal usage denoting taboo referent	3%	3%	hmmm, don't we all wish we had fucked that certain blonde in lourdes.....:-p // Oh n shaggin dead people // …that does not mean I am the village yokal idiot or that daddy fucked me with a rusty broken pitch fork…
Imagery based on literal meaning	-		[None]
Pronominal form with undefined referent	0%	1%	I Don't Care Who You Think You Are, Where You're From Or How Many Of You Think It'd Be Fun To Start Some Shit With Them, I Will Fuck You Up. Simple As That. // Occupation: No.1 Cunt Kicker-inner.
Personal: Personal insult referring to defined entity	27%	28%	tehe i am sorry.. i m such a sleep deprived twat alot of the time! lol // Maxy is the soundest cunt in the world!!!! // 3rd? i thought i was your main man number one? Fucker // write bak cunt xxx // You Godless bastard! // You evil cunt! Haha // bolllox you cunt! // what ya mean i'm no longer ur friend, ya bastard!! // CHEEKY LITTLE CUNT ! // lucky fuck
All	100%		

8. Discussion

Several factors affect the interpretations that can be placed upon the results. These are discussed below in order of increasing scope of interpretation.

First, to what extent does the measurement of swearing in a MySpace reflect the actual amount of swearing present? The measurements reported here are underestimates because of the restricted set of swear words used and in particular the omission of all reclaimed words, and the incompleteness of the search for non-standard spellings. There are very many non-standard spellings. For example, the following words starting with *shit* were too infrequent for inclusion in the analysis: *shitbag, shitball, shitee, shitest, shitface,*

shitfaced, shitgrind, shithole, shitin, shitless, shitload, shitsville, shitt, shittest, shittiest, shittin, shittiness, shitting, shittt, shittttt, shitttttttttttttttttt, shity, shityest, shitz, shitznitz, and *shitzooi*. Comparing figures 2 and 3, however, it seems likely that the standard swearing of fuck is so dominant that the overall results would not be significantly influenced by the inclusion of many more low frequency terms. The implicit assumption in the methods is that the effect of all the restrictions used will be similar for women and men, although this has not been verified.

Second, to what extent is the strength assessment of swear words accurate? The word frequency technique used is a limiting factor here because the strength of a word depends to some extent upon the context in which it is used. This particularly applies to abusive words that also have reclaimed usages that are much milder. Nevertheless, the words used in the main assessment seem to be unambiguous in their inherent swearing strength, although their perceived strength is likely to vary from community to community.

Third, to what extent does MySpace swearing reflect the owner's desired identity projection? This issue is discussed above but is at least a simplification and remains a method limitation.

Fourth, could MySpace swearing reflect offline conversational and other forms of swearing? This is problematic because MySpace users are not typical of the population of any country, for example probably tending to be younger and having regular access to the internet. There are also many likely differences between conversational and MySpace swearing, many along the lines of the seven swearing factors discussed above. For example, it seems likely that humorous swearing and unusual spellings are both relatively common. In contrast, expletives are probably rare and relatively trivial linguistic swearing forms may be less frequent than in conversation for those who swear continuously, simply as a time-saving exercise. The classification (Table 5) also suggests that playful insults and idiomatic or figurative uses are relatively common in MySpace and swearing for emphasis is relatively uncommon. Moreover, others have argued that web-based communication can be empowering for women (Herring 2003), which suggests that gender may affect the process of transferring between online and offline language. MySpaces could be seen as inherently mixed gender spaces, and therefore unlikely to contain communication about taboo topics of conversation that are normally restricted to same-gender groups, and seem likely to be gender-specific to some degree (e.g., Jay 1992:159). Nevertheless, is seems likely that the vocabulary of MySpace is often not too far divorced from that of informal offline communication in mixed gender groups in the extent to which swearing is considered acceptable, if not in the frequency and linguistic forms used. MySpace use is an important part of many people's lives, particularly for younger people, and so it seems likely that differences broadly reflect wider language use. In particular, it seems that the set of swear words used in MySpace would reflect those considered acceptable in informal conversations with offline friends.

Fifth, could MySpace swearing reflect the swearing of the wider population? A limitation for any such generalisation is the restricted user-base. This is particularly true in the U.K., where about 17% used social network sites in March-April 2007 (Dutton and Elsper 2007), although this conceals a larger proportion of younger users, for example 42% of students.

9. Conclusion

The results for UK MySpace members reveal more very strong and moderate swearing for males than for females, but a similar amount of strong swearing, with a suggestion of younger female MySpaces containing more strong swearing than male MySpaces. Very strong swear words are relatively rare, however, and so the most significant finding is that there is no gender difference overall for the use of stronger swear words (i.e., strong and very strong combined). The finding for moderate swear words is less robust because many moderate swear words were excluded (see Table 6). The results for US MySpace members reveal more strong swearing for males than for females, but a similar amount of moderate and very strong swearing. Given the relative rarity of very strong swear words, the most significant finding

seems to be that there *is* a gender difference overall for the use of stronger swear words (i.e., strong and very strong combined). The findings for moderate swear words are not robust, as for the UK case above. Moreover, the equality for very strong swear words is also not reliable because of the large gender difference in the very strong swear word *nigger* (and variations). Although predominantly used in a reclaimed context, *nigger* could still be seen as a strong or very strong swear word, unlike other swear words that become neutral when used in a reclaimed context (e.g., *gay*).

The results also showed that younger users had more swearing in their MySpaces than older users, with a disproportionate increase for females. Although not statistically significant, the amount of swearing between U.K. male and female MySpaces was closer for younger members. As identity projections in MySpace, this broadly matches previously established conversational norms, except for an absence of the U.K. gender gap for swearing. The figures reported are underestimates of the extent of strong swearing in MySpace because of the restriction of the main data to unambiguous swearing, and hence swearing must be seen as a normal rather than a deviant aspect of youth identity in MySpace.

The results can also be read as a partial reflection of current conversational swearing practices. As such, they are consistent with the swearing gender gap being strong in the U.S. but disappearing in the U.K., at least for strong swearing and younger users (which dominate MySpace). Nevertheless, because of differences between MySpace and spoken conversations it is not possible to draw reliable inferences about patterns of offline conversation, except perhaps concerning the extent to which a vocabulary is judged acceptable.

The U.K.-U.S. difference in gendered swearing is particularly significant because it is suggestive of a fundamental underlying difference in gender roles or expectations between the two countries. A possible cause or contributory factor is the rise in (Day, Gough, and McFadden 2004), and eventual acceptance of (Plant and Plant 2006:45), binge drinking amongst U.K. females. This has given rise to group terms such as *ladette culture* (Jackson 2006) which have replaced individualised descriptors of females with gender reversal behaviour such as *tomboy* or, more recently in the U.K., *geezer bird*. In the U.S. (and Mexico), recent research has cast considerable doubt on one linguistic gender stereotype: that women talk more than men (Mehl, Vazire, Ramirez-Esparza, Statcher, and Pennebaker 2007) but it seems that some important differences remain.

In terms of future research it would be interesting to compare MySpace swearing with conversational swearing, especially for younger users. In such cases the MySpace data could be used to provide initial hypothesis about the type of swearing used, including any new words developed. It would then also be useful to compare the results to identify types of swearing that is either specific to MySpace or to conversations. Also, more detailed analysis of terms of linguistic forms and other social factors (e.g., Mautner 2007) related to swearing would give a more comprehensive understanding of MySpace swearing, especially if combined with parallel offline investigations. Previous research has argued that adolescents are the linguistic leaders, rather than adults or children (Eckert 2003), and so it seems likely that gender equality in swearing or a reversal in gender patterns for strong swearing will slowly become more widespread, at least in social network sites. If so, this is an extremely important development for gender roles in the U.K., especially because swearing is closely related to psychological development and hence probably reflects much more fundamental shifts in the social psychology of the population.

Finally, an immediate and practical implication of this research is that because swearing amongst youth is normal rather than deviant behaviour, parents should not be shocked or concerned if their daughter or son has a MySpace containing strong swearing, although the reverse is perhaps not true (see Figure 3)!

10. Acknowledgements

Thank you to the referees and Graham Dwyer for very helpful comments and to Katie Horner for helping to check the data and for classifying the swear words.

11. Appendix

Table 6. Swear words considered and used, along with attributed strengths. Bold words were used in the main analysis.

Swear words	*Strength***
Cunt*, jew, **motherfuckin*, motherfucking*, muthafucker*, muthafuckin*, mutherfucker*,** nigga, niggah, niggas, niggaz, nigger, nigguh, paki	Very strong
Fuck*, fucked*, fucken*, fucker*, fuckin*, fucking*, fuckstick*, spastic	Strong
Arsehole, asshole, bastard, bollock, cock, dick, gay, **piss*, pissin*, pissing*,** poof, poofter, poofy, prick, pussy, queer, shag, **shagged, shagging, twat, wank, wanker, wanking,** whore	Moderate
Arse, arsed, ass, bitch, bugger, butthole, christ, cow, dickhead, dipshit, fanny, fart, jesus, moron, pissed*, retard, screw, screwed, screwing, shit*, shite*, shithead*, shittin*, shitty*, slag, slagged, slut, tit*, titties*, tosser	Mild
Bap, bimbo, bird, bloody, bonk, bonking, boob, bullshit*, butt, butthead, crap, damn, dork, dorky, git, god, hell, hussy, idiot, jerk, jug, knocker, pig, pillock, pimp, sod, tart, tarty, turd, wuss	Very mild

*used in the U.S. word list

**words are not categorised by the degree of offence given but by public average perceptions, which depend partly upon usage. For example *nigger* and *queer* could be extremely offensive in an abusive context but inoffensive in a reclaimed context. Similarly, *Jew* and *gay* could be used as insults or as neutral self-descriptions.

12. References

Abrahams, R.D. 1962. 'Playing the dozens', *Journal of American Folklore* 75 (297), pp 209-220.

Andersson, L.-G., and P. Trudgill. 1992. *Bad language*. London: Penguin.

Bailey, L. A., and L. Timm. 1976. 'More on women's - and men's – expletives', *Anthropological Linguistics* 18 (9), pp 438-449.

Bayard, D., and S. Krishnayya. 2001. 'Gender, expletive use, and context: Male and female expletive use in structured and unstructured conversation among New Zealand university students', *Women and Language* 24 (1):1-15.

Bell, S.E., and S.M. Reverby. 2005. 'Vaginal politics: Tensions and possibilities in The Vagina Monologues', *Women's Studies International Forum* 28 (5), pp 430-444.

Benwell, B. 2001. 'Male gossip and language play in the letters pages of men's lifestyle magazines', *The Journal of Popular Culture* 34 (4):19-33.

Blair, David. 2006. *Wittgenstein, language and information: "Back to the rough ground!"* Amsterdam: Kluwer.

boyd, d.. 2006. 'Friends, Friendsters, and MySpace Top 8: Writing community into being on social network sites', *First Monday* 11 (2), Retrieved June 23, 2007 from: http://www.firstmonday.org/issues/issue11_12/boyd/index.html.

boyd, d.. in press. 'Why youth (heart) social network sites: The role of networked publics in teenage social life', In *MacArthur Foundation on Digital Learning, Identity Volume*, edited by D. Buckingham. Cambridge, MA: MIT Press.

boyd, danah, and Jeffrey Heer. 2006. 'Profiles as conversation: Networked identity performance on Friendster', *Proceedings of the Hawai'i International Conference on System Sciences (HICSS-39, January 4-7)*, Retrieved July 3, 2007 from: http://www.danah.org/papers/HICSS2006.pdf.

Burnard, L. 1995. *Users' reference guide to the British National Corpus*. Oxford: Oxford University Computing Services.

Coates, J. 2003. *Men talk*. Maldon, MA: Blackwell.

Day, K., B. Gough, and M. McFadden. 2004. 'Warning! alcohol can seriously damage your feminine health', *Feminist Media Studies* 4 (2), pp 165-183.

Dewaele, J-M. 2004. 'The emotional force of swearwords and taboo words in the speech of multilinguals', *Journal of Multilingual and Multicultural Development* 25 (2&3):204-222.

Dooling, R. 1996. *Blue streak: Swearing, free speech and sexual harassment*. London: Random House.

Dutton, W.H., and E.J. Elsper. 2007. *The Internet in Britain 2007*. Oxford: Oxford Internet Institute.

Eckert, P. 2003. 'Language and gender in adolescence', In *The Handbook of Language and Gender*, edited by J. Holmes and M. Meyerhoff. Oxford: Backwell.

Ellison, N, C Steinfield, and C Lampe. 2006. 'Spatially bounded online social networks and social capital: The role of Facebook', *Annual Conference of the International Communication Association*:Retrieved April 23, 2007 from: http://msu.edu/~nellison/Facebook_ICA_2006.pdf.

Fono, D., and K. Raynes-Goldie. 2005. 'Hyperfriendship and beyond: Friendship and social norms on Livejournal', Association of Internet Researchers (AOIR-6), Chicago.

Golder, S. A., D. Wilkinson, and B. A. Huberman. 2007. 'Rhythms of social interaction: Messaging within a massive online network', 3rd International Conference on Communities and Technologies (CT2007), East Lansing, MI.

Green, J.A. 2003. 'The writing on the stall', *Journal of Language and Social Psychology* 22 (3):282-296.

Grimm, M. 2003. 'When the sh*t hits the fan',. *American Demographics* 25 (10):74-75.

Herring, S. C. 2003. 'Gender and power in on-line communication', In *The Handbook of Language and Gender*, edited by J. Holmes and M. Meyerhoff. Oxford: Backwell.

Hey, V. 1997. *The company she keeps: An ethnography of girls' friendships*. Buckingham: Open University Press.

Hinduja, S., and J.W. Patchin. In press. 'Personal information of adolescents on the Internet: A quantitative content analysis of MySpace', *Journal of Adolescence*.

Hughes, G. 1991. *Swearing: A social history of foul language, oaths, and profanity in English*. Oxford: Blackwell.

Jackson, C. 2006. ''Wild' girls? An exploration of 'ladette' cultures in secondary schools', *Gender and Education* 18 (4), pp 339-360.

Jay, T. 1980a. Sex roles and dirty word usage: A review of the literature and reply to Haas. Psychological Bulletin, 88(3), pp. 614-621.

Jay, T. 1980b. 'A frequency count of college and elementary school students' colloquial English'. *Catalogue of Selected Documents in Psychology*, 10, 1.

Jay, T. 1992. *Cursing in America: A psycholinguistic study of dirty language in the courts, in the movies, in the schoolyards and on the streets*. Amsterdam: John Benjamins.

Jay, T. 2000. *Why we curse: A neuro-psycho-social theory of speech*. Amsterdam: John Benjamins.

Jay, T., K. King, and T. Duncan. 2006. 'Memories of punishment for cursing', *Sex Roles* 55, pp 123-133.

Liladhar, J. 2000. 'J. Eclair, 'The rotting old whore of comedy': A feminist discussion of the politics of stand-up comedy', *Working Papers on the Web* 1 (1), Retrieved September 12, 2007 from: http://extra.shu.ac.uk/wpw/femprac/liladhar.htm.

Lloyd, M. 2007. 'Rear gunners and troubled privates: Wordplay in a dick joke competition', *Journal of Sociolinguistics* 11 (1) pp. 5-23.

Marsh, J. 1998. *Word crimes: Blasphemy, culture, and literature in nineteenth-century England*. Chicago: University of Chicago Press.

Mautner, G. 2007. Mining large corpora for social information: The case of elderly. *Language in Society* 36 (1):51-72.

McEnery, T. 2005. *Swearing in English*. London: Routledge.

McEnery, T, and Z. Xiao. 2004. 'Swearing in modern British English: the case of fuck in the BNC', *Language and Literature* 13 (3), pp 235-268.

Mehl, M., and J. Pennebaker, 2003. The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology, 84 (4)* pp. 857-870.

Mehl, M., S. Vazire, N. Ramirez-Esparza, R. Statcher, and J. Pennebaker, 2007. 'Are women really more talkative than men?' *Science, 317(5834),* pp. 82-82.

Meyer, C., R. Grabowski, H.-Y. Han, K. Mantzouranis, and S. Moses, 2003. 'The world wide web as linguistic corpus'. *Language and Computers, 46*(1), pp. 241-254.

Mills, J. 1991. *Womanwords*. London: Virago.

Montagu, A. 1967. *The anatomy of swearing*. Basingstoke: MacMillan.

Ooi, V.B.Y. 2000. 'Aspects of computer-mediated communication for research in corpus linguistics' In: *Language and Computers, New Frontiers of Corpus Research. Papers from the Twenty First International Conference on English Language Research on Computerized Corpora Sydney 2000*, pp. 91-104.

Payne, J.S. 2005. Synchronous CMC, working memory, and L2 oral proficiency development *Language Learning & Technology* 9 (3), pp. 35-54.

Plant, M., and M. Plant. 2006. *Binge Britain: Alcohol and the national response*. Oxford: Oxford University Press.

Prescott, L. 2007. Hitwise US consumer generated media report. Retrieved March 19, 2007 from: http://www.hitwise.com/.

Rayson, P., G Leech, and M. Hodges. 1997. 'Social differentiation in the use of English vocabulary', *International Journal of Corpus Linguistics* 2 (1), pp 133-152.

Sapolsky, B S., and B K. Kaye. 2005. 'The use of offensive language by men and women in prime time television entertainment', *Atlantic Journal of Communication* 13 (4) pp. 292-303.

Schapiro, N. A. 2002. ''Dude, you don't have Tourette's' Tourette's Syndrome, beyond the tics', *Pediatric Nursing* 28 (3):243-253.

Sigel, L. Z. 2000. 'Name your pleasure: The transformation of sexual language in nineteenth-century British pornography', *Journal of the History of Sexuality* 9 (4), pp 395-419.

Smith, S. 1998. 'The social meanings of swearing: Workers and bad language in late imperial and early soviet Russia', *Past and Present* 160, pp 167-202.

Stapleton, Karen. 2003. 'Gender and swearing: A community practice', *Women and Language* 26 (2):22-33.

Stokoe, E, and D. Edwards. 2007. ''Black this, black that': racial insults and reported speech in neighbour complaints and police interrogations', *Discourse & Society* 18 (3), pp 337-372.

Thelwall, M. submitted. 'Social networks, gender and friending: An analysis of MySpace member profiles', Retrieved August 23, 2007 from: http://www.scit.wlv.ac.uk/~cm1993/papers/MySpace_d.doc.

Thelwall, M., and L. Price. 2006. 'Language evolution and the spread of ideas: A procedure for identifying emergent hybrid word family members', *Journal of the American Society for Information Science and Technology* 57 (10), pp 1326-1337.

Trudgill, P. 1972. 'Sex, covert prestige and linguistic change in the urban British English of Norwich', *Language in Society* 1 (2), pp 179-195.

van Lancker, D., and J. L. Cummings. 1999. 'Expletives: Neurolinguistic and neurobehavioral perspectives on swearing', *Brain Research Reviews* 31 (1), pp 83-104.