

‘Am I normal?’ Teenagers, sexual health and the internet

Kevin James Harvey^{a,*}, Brian Brown^b, Paul Crawford^c,
Aidan Macfarlane^c, Ann McPherson^d

^a*School of English Studies, University of Nottingham, University Park, University Boulevard, Nottingham, Nottinghamshire NG7 2RD, UK*

^b*De Montfort University, UK*

^c*Independent International Consultant in the Strategic Planning of Child and Adolescent Health Services, Oxfordshire, UK*

^d*University of Oxford, UK*

^e*School of Nursing, Queen’s Medical Centre, University of Nottingham NG7 2UH, UK*

Available online 14 May 2007

Abstract

The development of new communication technologies has created a wide variety of new fields in which human beings can construct identities. The past decade has seen a proliferation of opportunities to use the internet for health related advice and information and many new sites have been created where participants can construct identities, formulate problems and seek solutions concerning health related issues. This paper will report on a study of emails written in to a UK-based website concerned with health issues for young people. Our analysis was driven by corpus linguistics, a computational methodology for interrogating extensive datasets, and we have combined both quantitative and qualitative approaches to the study of language. We interrogated a 400,000 word dataset of messages and were able to identify terms whose usage was elevated compared to the English language as a whole. As well as personal pronouns, these included many terms related to sexual health and bodily development, as well as terms such as ‘normal’ and ‘worried’ which were identified for further exploration. Whereas previous research on sexual health has discovered the use of vague terms and euphemisms, here, young people described themselves, their anatomy and their identities in meticulous detail. This study enables us to define the role of health topics raised, the presentation of health concerns, and contributes towards the discovery of a distinctive ‘genre’ of health messages concerning sexual health which differs from that found by other researchers concentrating on face to face encounters. In conclusion we suggest that for researchers and practitioners working in health with young people in the medium of English there could be valuable lessons in communication to be learned from examination of corpora of the health care language concerned.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Adolescents; Internet and electronic communication; Corpus linguistics; Sexual health; Email; UK

Introduction

The rise of the internet has led to a substantial increase in the number of websites dedicated to providing health advice and information. Many new fora have been created whereby internet users can formulate problems and seek solutions concerning

*Corresponding author. Tel.: +44 7913703451.

E-mail addresses: kevin.harvey@nottingham.ac.uk (K.J. Harvey), brown@dmu.ac.uk (B. Brown), paul.crawford@nottingham.ac.uk (P. Crawford), Aidanmacfa@aol.com (A. Macfarlane), ann.mcpherson@dphpc.ox.ac.uk (A. McPherson).

health related issues. Although many commentators have reported that the electronic landscape is populated by many unreliable, misleading and unscrupulous sites purporting to provide accurate medical information (e.g. Barkham, 2000; Henk, 2002; Thurlow, Lengel & Tomic, 2004), the internet nevertheless remains a popular source of medical advice and information with the potential to contribute to positive health care outcomes and intervention (Car & Sheikh, 2004; Gray, Klein & Noyce, 2005).

With the rise of internet health resources over the last decade there has been an increased demand for health provision tailored to the needs of young people (Aynsley-Green et al., 2000; Jones, Finlay, Simpson, & Kreitman, 1997), resulting in a growth of internet-based resources tailored to delivering reliable and accurate health information for young people. These sites typically provide interactive advice in accessible, non-technical language through which young people can freely express their health questions to professionals, as well as sharing their concerns with fellow teenagers. Consequently the internet has become a popular source of health advice and information for adolescents, with the electronic gateway offering confidential advice and information that might be otherwise hard or compromising to obtain (Borzekowski & Rickert, 2001; Suzuki & Calzo, 2004). For example, the UK based Samaritans charity, an organization that provides confidential emotional support in response to enquiries by distressed individuals, report that since their email service was introduced in 1995, 'it has been used by many young people as their preferred way of discussing difficult feelings', noting that this service had increased by almost 80% between 2002 and 2004 (Samaritans, 2004).

The issues raised by researchers on internet health in different countries are intriguingly similar, especially given the differentiation in face to face health care cross-nationally. Nations with substantial public provision yield similar concerns to those with largely insurance-funded systems. On both sides of the Atlantic, and even as far afield as China, the anonymity afforded by the internet is valued (Lou, Zhao, Gao, & Shah, 2006; Rice, 2006; Rogers & Mead, 2004) and this is particularly important for young people, since they are often reluctant to disclose potentially embarrassing and sensitive problems and request personal health advice from their doctors (Ackard & Neumark-Sztainer, 2001; Biddle, Gunnell, Sharp, & Donovan, 2004; Klein &

Wilson, 2002). They make fewer visits to office-based doctors than other age groups (Monson, Jackson & Livingston, 1996). This is of concern given that adolescence is a time of physical, emotional and social change with profound needs to address numerous delicate questions concerning health, sexuality, and relationships (Klein & Wilson, 2002; Suzuki & Calzo, 2004).

While eager to broach sensitive concerns about their health with other people, adolescents often find it difficult to articulate their concerns to adults (Boekeloo, Schamus, Cheng, & Simmens, 1996). Many adolescents want to discuss health concerns with their GPs but are very particular about whom they consult, being only likely to see a practitioner with whom they feel comfortable (Kapphahn, Wilson, & Klein, 1999). According to Holland, Ramazanoglu, Sharpe and Thomson (1996), teenagers are often reluctant to talk to parents about sex and sexual health because of potential embarrassment and in case their parents think they are sexually active.

With such an apparent generational gap, it is important to advance understandings of how adolescents request advice on sensitive matters such as sexual health so as to better inform the strategies of engagement with this population by health practitioners. In this context we would argue that corpus linguistics is a valuable methodology for studying adolescent health communication, providing not only an original and valid means for identifying the issues and troubles young people face in their everyday lives (Michaud, Suris, McPherson, & Macfarlane, 2004) but also as a means of describing a distinctive 'genre' of health messages concerning sexual health which differs from that of much other research which has focussed on face to face encounters.

Corpus linguistics as a means for interrogating health communication

Widely used in other areas of linguistic analysis, corpus linguistics is a relatively new application in relation to health care communication. However, its benefits are quickly coming to be recognised by medical researchers and professionals and recently there has been a variety of insightful corpus studies into the domain of health care discourse (e.g. Adolphs, Brown, Carter, Crawford, & Sahota, 2004; Brown, Crawford, & Carter, 2006; Skelton & Hobbs, 1999; Skelton, Wearn, & Hobbs, 2002;

Thomas & Wilson, 1996). Studies in health communication have almost exclusively utilised observational, qualitative methods (Skelton & Hobbs, 1999). Such research adopts a close language-based focus in order to provide a penetrating description of the medical interaction/practice under consideration. Yet qualitative studies in health communication often work with small data sets. Critics allege that a drawback of using limited samples of data, such as a small set of doctor–patient encounters, is that conclusions will not necessarily represent the wider field of language used in that particular domain (Stubbs, 1997). By contrast, corpus linguistic studies involve interrogating large sets of data so as to describe and explicate patterns of communication. The methodological advantage of analysing large data sets is that they are more likely to be representative of the text type under consideration than short fragments of data. Using large data sets, furthermore, allows the analyst to account for a wide range of variation which might be present in the texts and therefore ground generalisations on more substantial and representative textual evidence.

A distinguishing feature of corpus linguistics is its use of computer software to generate word frequencies and concordances. The calculation of lexical frequencies is an important part of linguistic description, since it allows precise quantification of the most commonly used words in a given communicative context. By identifying strings of text containing such frequently or commonly occurring words, concordance lines provide a picture of how these important items are being used in a particular communicative context. The combined computational tools (word frequency lists and concordance outputs) therefore harness both quantitative and qualitative methods and provide a novel and increasingly popular approach to the analysis of language and communication in health settings.

Method

We interrogated the content of a 400,000 word corpus of emails sent to the popular adolescent health website, Teenage Health Freak (www.teenagehealthfreak.org). Operated by two doctors based in the UK and specialising in adolescent health, Teenage Health Freak has been running and continuously updated on a weekly basis since 2000. The site is designed to be user friendly, interactive, confidential and evidence based, em-

ploying non-technical, accessible language. During the design of the site, young people themselves were actively involved in its construction in order to incorporate their views on what kind of services and facilities would be beneficial to them.

Between February 2000 (the year of its launch) and January 2007 the Teenage Health Freak website has received 60,000,000 visits, an average of 52,864 hits a day. The website receives and responds to emails that are communicated anonymously and in confidence by adolescents seeking health advice and information. Since its launch the site has received over 120,000 emails from young people. While we cannot make precise demographic claims about adolescent health or the nature of teenagers' concerns in their entirety from this information, the experience it represents is not inconsiderable and merits inquiry in its own right in light of the growing role of interactive internet-based health advice. Although the doctors do not respond to all of the questions, given the large amount of messages sent to the site, new key questions and answers appear on the site daily. Unlike data from other interactive health websites which, in sections such as FAQs, reformulate and standardise messages sent by users seeking health advice, the data used in this study reproduce the entire, unedited queries and concerns posted to the site doctors. Consequently this allows for access to the original word choice and syntactical construction of the email requests, retaining any grammatical or orthographical inconsistencies and all the nuances of individual expression.

We were given permission by the Teenage Health Freak operators to collect and analyse the emails sent to the website. We collected data sent to the website from January to August 2004. This amount, covering over half of 1 year's worth of emails, provides a substantial snapshot of the health concerns articulated daily by teenage contributors. The Teenage Health Freak website possesses and displays a privacy policy that informs potential senders of requests for advice that the information communicated in their requests may be used for research purposes and that, in using the website to transmit such information, they consent to the collection and use of the data which they supply.

In order to conduct our initial computational analysis of the email corpus, we employed the WordSmith Tools software program (Scott, 1996). Once the pattern of word frequencies had been used to establish the 'genre' or 'register' of

communication in this context, it was possible to adopt the recommendations of Orpin (2005) and Conrad (2002) to follow up particular terms in a more fine grained fashion. This part of the analysis enabled us to identify the kinds of problems which were being formulated, drawing upon the notions of discourse analysis. For example, attention was paid to the collocation of terms—that is, as we shall see, some parts of the body were routinely associated with problems of both a physical and ethical nature. Equally, ideas of how language can signal the writers' 'stance' (Precht, 2003) or 'appraisal' (Mack-en-Horarik, 2003) were used to illuminate the implied evaluations embedded in the use of certain terms.

Results and discussion

The entry point into our collection of adolescent health emails involved using corpus linguistics software to compare this dataset with a corpus of general English. This yielded a number of keywords (see Table 1 below) that appear more than one would expect in the English language as a whole, based on a comparison with the cancode corpus, a 5 million word database of general english. Table 1 lists the absolute frequencies of the keywords salient to the teenage health messages and their comparative frequencies in the cancode corpus. As well as expressing both sets of frequencies as percentages of the total word counts, the range of keywords provides an indication of the central health themes and verbal choices appearing in the adolescent corpus and is thus able to identify patterns of communicative style that are unique to this particular health language variety.

Unsurprisingly, in the teenage email corpus first person pronouns and connected terms like 'am' are very frequently observed. The high occurrence of the present tense first-person singular form of the auxiliary verb 'be' ('am' and its contracted form 'I'm') anchors health-related concerns in the present, signalling their currency at the time of articulation: 'I am worried about...' What is particularly interesting is the frequency of terms related to sexual health, which significantly exceed the frequency we would expect in English as a whole. Although the website solicits input across a whole range of health issues, it is sexuality and reproductive health that predominates in the emails, inasmuch as terms like 'penis', 'pregnant', 'period' and 'gay' feature far more frequently than they do

Table 1
Frequencies of keywords in the 'teenage email' corpus

WORD	TEEN FREQ.	%	CANC FREQ.	%
1 MY	9775	2.55	24,435	0.47
2 I	25,287	6.59	167,377	3.24
3 AM	3594	0.94	4691	0.09
4 SEX	3208	0.84	3421	0.07
5 IM	2234	0.58	2265	0.04
6 ME	4659	1.21	19,068	0.37
7 PENIS	1480	0.39	1508	0.03
8 HELP	1834	0.48	3163	0.06
9 QUIZ	1273	0.33	1293	0.03
10 ANN	1184	0.31	1335	0.03
11 DONT	1110	0.29	1110	0.02
12 ASKED	1443	0.38	2353	0.05
13 PREGNANT	1092	0.28	1214	0.02
14 QUESTION	1374	0.36	2325	0.04
15 HAVE	6237	1.63	39,324	0.76
16 IS	6924	1.81	47,740	0.92
17 DO	5655	1.47	36,146	0.7
18 BOYFRIEND	848	0.22	965	0.02
19 BULLYING	785	0.2	793	0.02
20 PERIOD	884	0.23	1155	0.02
21 DR	821	0.21	958	0.02
22 PLEASE	1265	0.33	2943	0.06
23 DRUGS	757	0.2	980	0.02
24 HOW	2377	0.62	11,492	0.22
25 WORRIED	744	0.19	1103	0.02
26 U	743	0.19	1193	0.02
27 GAY	616	0.16	727	0.01
28 NORMAL	746	0.19	1291	0.02

Keywords based on a comparison of a 400,000 word subsample of adolescent health emails and the 5 million word CANCODE corpus.

in general English. A few other oddities relate to the website itself. 'Ann' corresponds to the virtual character 'Dr. Ann', one of the Teenage Health Freak's resident doctors and adolescent health specialist who routinely responds to the queries submitted by teenagers, and 'quiz' corresponds to the quizzes the website presents as means of providing educational material: 'So you think you know about AIDS?' and so on.

The analysis of corpus material allows a process of 'drilling down' to address progressively more detailed and contextualised levels of analysis once the key features of the communicative terrain have been apprehended. Consequently, we can examine some of the occurrences of terms relating to anatomy and sexuality *in situ* where correspondents to the website describe difficulties and concerns in detail. If we take for example the term 'penis'—indicated as being overrepresented in Table 1—it is

possible to examine some of the contexts in which the term occurs:

- I have already ripped my foreskin when I 1st a good while ago but recently when having sex Ive ripped it further and is very painful to have sex it happened 4 months ago what can I do as Im embarrassed to go to the doctor?
- I was having sex when my boyfriend told me to suck his penis or he would tell his freinds that I open and he was leaving me. Desperate girl.
- Dr. Ann I have a 4 inch penis and the boys in my class all say they have 6 or 7 inches im I normal?

This small selection should suffice to demonstrate the level of detail in which anatomies and the social practices within which they were embedded are described. This contrasts with the degree of vagueness or apparent difficulty in calling experiences to mind that practitioners and researchers have found elsewhere (e.g. Mason, 2005; Stewart, 2005; Weijts, Houtkoop, & Mullen, 1993). For example, advice workers find that young people are apt to say they ‘don’t know’ whether they have had sex, or speak ambiguously concerning their activities (Mason, 2005). Here, by contrast, not only is there considerable detail but the concerns of the youngsters are ethically and morally situated, inasmuch as their accounts are formulated so as to identify that something is wrong with the physical integrity, feeling, size or relational activity surrounding the penis.

The detail of these complaints was often ‘atopic’ in that they did not contain an explicit contrast with what the writers thought *should* be the case. This has also been noted in anthropological accounts of complaint narratives (Lee, 2003) and in conversation analytic accounts of complainable matters (Drew, 1998). The unfairness of the situation is implied without describing any alternative which would be fair. This is a commonplace motif in tales of injustice and confirms a kind of solidarity between speakers and hearers. When the problems are re-told they are even more effective if the departure from fairness is left unsaid; e.g. that one’s penis should be pain free or of a size claimed by other people and that one’s boyfriend should not threaten to damage one’s reputation among the peer group. ‘Normative mundanity’ is typically ‘unmarked and unremarkable’ (Bostock, 2002, p. 352).

Either implicitly or explicitly, a ‘normative structure’ is apparent in many of the messages. This

can be conceptualised as the ‘structure of meaning-in-use’ (Milliken, 1999, p. 132) that ‘frames’ decisions and evaluations (Keck & Sikkink, 1998; Payne, 2001) and helps set a framework of reference for decisions about states of health and the suitability of medical intervention or advice. Often, this centred upon overarching constructs concerned with whether a situation or symptom was ‘normal’ or with being ‘worried’ about an issue.

Let us take as an example the word ‘normal’, which appears more frequently in the present corpus than in everyday English as a whole. Some examples of how the authors of the messages apply the keyword ‘normal’ are given in the concordance lines below in Table 2.

The concordance lines generated by the software allow us to isolate all the instances of the node word ‘normal’—which are, for ease of linguistic interrogation, vertically stacked upon one another—and thereby examine the communicative context in which they appear. Grammatically, ‘normal’ performs a range of expressive functions, modifying and referring to specific health concerns and anxieties (e.g. ‘normal weight’, ‘normal size’, ‘is this normal’, ‘is it normal’), while being employed in a more general way to convey a sense of overall normality: ‘am I normal’—against which the placing of specific problems are brought into textual relief. Thus the various grammatical realisations of ‘normal’ can be seen to relate to a range of issues, including mood, sex or age related norms, such as when it is usual to begin one’s periods or how regular they should be, presumably when one is in the early stages of maturity. In line with these physical, emotional, and psychological changes (Suzuki & Calzo, 2004), the focus on, and concern with, what is deemed to be normal can be seen as a reflection of the developmental stage of adolescence, particularly early adolescence, during which time younger teenagers are more occupied with personal identity than older teenagers who are more likely to have resolved some of their most urgent identity issues (Erikson, 1959; Subrahmanyam, Greenfield, & Smahel, 2006). Furthermore, it is possible that these questions addressed to the online practitioner concerning whether a given complaint is normal or not bear out the finding that teenagers use the internet as means of cross-validating opinions and advice they receive from personal sources (Gray, Klein, Noyce, Sesselberg, & Cantril, 2005).

Yet, equally, as the concordance lines above reveal, we can see how the concept of normality is

Table 2
20 Randomly selected concordance lines of ‘normal’

I'm 12, I'm 5, 3 ft and 42 kg is this a	normal	weight or is it too light?
or being flat chested. Worried boobs aren't	normal	size. episodes of Bulimia
seen that is normal. But I dont want to b	normal	I want to be thin. I find it insulting
14 and I havent started my period am I	normal	I masterbate. I use
sick for no apparant reason. Is it	normal	to miss a period for 3 months
thinking about becoming a transexual. Is it	normal	to do this?
I still have a lot of discharge is this	normal	??
been a little depressed recently. Is this	normal	?
ward and I havent got any pubic hair am I	normal	? I am 13 and my name
this white stuff in my under wear, am I	normal	or do I have a disease.
during intercourse how far back does the foreskin go?		
Mine goes over the head of the penis but no further. Is		
this	normal	Will I be able to have intercourse.
I have a 4 inch penis and the boys in my class all say they		
have 6 or 7 in am I	normal	?
Is it	normal	for my balls not to have dropped by this time in life
Is there such a thing as pubic dandruff? How would u get		
rid of it? Is it	normal	t cut ur pubic hair?
Is it	normal	before your period to have a pain in only one breast, only
		round the nipple area? could that mean I might be
		pregnant even if I haven't fooled with a guy?
I can feel a lump inside my vagina, about 4 cm in, it isnt		
tender or painful to touch, but is it	normal	and if my boyfriend feels it will he be freaked out?
is phone sex	normal	because I've had it a couple of times with 2 boys
I've been sleeping with my boyfriend for 3 month now		
but am stil bleeding after sex sometimes heavily, is this	normal	or should I see a doctor?
I had my first period in December but the I skipped		
January is this	normal	

loaded with other meanings and positive valuations. That is, to someone concerned about secretions visible in their underwear, for example, normal might connote an imagined normative cleanliness and dryness whereas to someone concerned about the onset of periods the normativity might refer to the desirability of not being left out when one's peers begin menstruating. If we take a more fully developed picture of the kinds of contexts in which normal occurred, it is more obvious what it is being accomplishing in the correspondents' questions and how it contributes to a 'normative structure' (Caboni et al., 2005) or indeed a 'structure of feeling' (Ang, 1985) among the correspondents to the website. A further selection of 'normal' questions is presented in Table 3.

Looking at the way normality and its opposites are formulated in these questions, it appears that it is often part of a contrastive statement, such that the event or phenomenon is described and then an invited contrast or comparison is posed: 'is this normal?' It may not be speculating too much to see the term 'normal' as one which indicates stance (Precht, 2003) and appraisal (Macken-Horarik,

2003). That is, when people say 'normal' in this corpus it is often in the context of something they think is wrong within a normative structure, rather than in the sense of being merely curious as to whether things are statistically usual.

As Link (2004) points out, the notion of normality presupposes a particular kind of individual but also a particular kind of society—one in which data can be collected so as to establish those very norms against which the individual is judged. Moreover, the form of these questions and the answers they might receive reflects what Link (and Foucault (1990) before him) has referred to as the pedagogisation of sex. That is, it has become a matter for the giving and receiving of instruction, information and advice. Rather than the pattern of pervious epochs where it was undertaken in a more haphazard fashion, it is presently a mappable, knowable territory susceptible to expertise, direction and tutelage. Thus sexuality and sexual health become matters which can become assembled to the self, and offer, as Gergen (1991, p. 69) says '...the acquisition of multiple and disparate potentials for being'. The reconciliation and consolidation of

Table 3
Questions from the teenage health corpus using the word ‘normal’

-
- I am addicted to cerial. Is this normal?
 - Hi only one of my balls have dropped, is this normal? will the other one drop in time?
 - Dear Dr. Ann, My cum is not white, its clear with some white in it. Do I have a problem? Is this normal? Whta can I do to change this?
 - Hi Dr A. I fink iv started my periods but im not sure coz it woz brown. I told my mum and she sed I had started and it was normal 4 my fist period. But Im stil worried coz sumtimes ders brown stuf and sumtimes ders nowt der. Plz help
 - Hey I am a 16 female, When me and my boyfriend have sex I sometimes fart at of my viginia, is this normal?
 - Right by the hole in my vagina, there is. . . a thing like a ball. Is this my g-spot or is this normal??
 - Dr. Ann, I have a question about my inner lips in my vagina one inner lip is small and the other inner lip is bigger and it kinda sticks out what do I do and is that normal?
-

coherent personal narratives for both normative and minority sexualities has been identified as playing an important role in the development of sexual identities in both interpersonal and clinical contexts (Cohler & Hammack, 2007; DeSocio, 2005; Plummer, 1995).

Let us take another example of a relatively high frequency term in Table 1: ‘worried’. This too is suggestive of a particular kind of normative structure or kind of judgement. An examination of the corpus tells us what sorts of things people claim to be worried about. A small selection of the kinds of problems yielding usage of the term worried is presented in Table 4.

Here we can see a variety of expressive and rhetorical strategies used by adolescents in order to communicate the uniqueness (to them) of their concerns. As with the use of ‘normal’ in Tables 2 and 3 above, ‘worried’ is utilised in a range of collocational constructions, appearing in conjunction with specific issues (‘I am worried I may be pregnant’) as well as being employed to articulate a more general, yet no less emphatic, sense of distress: ‘I’m really worried’. Also, in the first example above, the import of the troubling experience articulated by the sender is marked and reinforced by the use of repetition, a common rhetorical strategy in the communication of emphasis and persuasion (McCarthy & Carter, 1994). Unlike classical formulations of the sick role, where social actors conjointly recognise and accept the sufferer’s indisposition, worry was here often accompanied by

Table 4
‘Worried’ narratives

-
- I am worried I may be pregnant but am still having normal regular menstrual cycles. I have taken a couple of home pregnancy tests and they say negative but Im worried about the time I do the tests as they say you should take it on the first day of your missed period, so if Im still having my periods, whens the best time to take the test?
 - I had protected sex 24h ago the condom did not break I am still worried that it may leaked. I was wondering how I would be able to make sure I am not pregant and a couple hrs l8er after sex I recieved my period can I get pregant like that?
 - I started my period roughly two years ago. It’s been quite regular, but all of a sudden, I’ve missed about 4 periods and still haven’t come on. I’ve never had sex though! I’m really worried. Why is this and what do I do?
 - I have been worried about what is wrong with my body. My Period was never really reglur ever sinceI started having it. Now I have been sexually active in the past 2 years and I haven’t had my period since last May. I’m scared to go to the doctor. What should I do???
 - Hi ya ann this is really imbarresing for me but here we go well u c I keep getting this really smelley discharge and its ok some days and bad on others specially leading up to my period I havent had sex or anything like that and Im worried on asking my mum and a doctor Im really worried can I do anything to stop this plzs plzs help I no you are very bissy but I need some help plzs thank you ever so thanks again yours in need ps any good advice on spots and when u get sweaty at school thanks;
 - I’ve had my period since I was about 12 and a half but they’re still very irregular, should I be worried by this? Also, I get vaginal discharge, all the time, Is this normal? Or should I only get it at certain times of the month?
 - Hiya doctor ann I had a really short and light period on holiday then a few days after it had finishd I got this reeli thick brown sticky discharge and ive had it for about 2 weeks now its not itchy or anything but I sometimes get cramps, please reply soon Im pretty worried n dont tell me 2 go to the doctors coz I wont thanx x x x x
-

a desire for secrecy: ‘I dont want to tell my mum I had sex ...cos she would go mad!’.

Many of the correspondents seemed to be experiencing worry where the possibility of pregnancy or menstrual disruption were concerned, judging by the collocations here and in the data set as a whole. It is as if underlying this pattern there is a substantive expectation that the cycle will be regular and departure from this occasions concern. Of course many people experience menstrual irregularity, and later in life this may not be identified as a problem but in one’s early experiences of menstruation, regularity may be perceived as normative. This then is reminiscent of Harre’s (1986) and Bedford’s (1957) idea that emotions are a kind of value judgement. This highlights another

use of corpora such as this in exploring lay theories and folk beliefs about health.

More broadly it is possible to see these ideas about worry, and the hoped for resolution through advice and instruction, as part of a “tutelary complex” (Donzelot, 1979). This term encompasses the complex of institutions designed to shape subjectivity which might include education, health and, perhaps latterly, advice available over the internet. Recent work by other scholars on everyday explanations for illness has disclosed a rich variety of lay explanatory frameworks. For example, Popay et al. (2003, pp. 7–9) discovered explanations for ill health ranging from ‘beer fags egg and chips’ to ‘worry and stress’ and which included ‘worse housing, high unemployment and a lack of hope in the area’. These kinds of descriptions and explanations of what they are worried about connect to implicit theories concerning personal estimations of risk (French, Sutton, Marteau, & Kinmonth, 2004).

The formulations of worry identifiable here have implications for the growing research interest in ‘worry’ on the part of health professionals (Laakso, Niemi, Grönroos, Aalto, & Karlsson, 2005) and students of emotion, cognition, anxiety and stress (Brosschota, Gerinb, & Thayerc, 2006). Borkovec, Ray, and Stfber (1998) introduced a working definition of worry as a chain of negatively affect laden, relatively uncontrollable thoughts and images. It represents engagement in mental problem solving on an issue whose future outcome is uncertain but one or more negative outcomes is possible. By contrast, empirical examinations of the occurrence of worry suggest communicative, performative, and normative aspects. Thus, research of the kind we have described here embeds the notion of worry within the shared communicative encounter between health care provider and recipient and helps to expose some of the precursors such as normative expectations about bodily functions, comparative judgements and the form and content of impending subjective risks. This might also correspond to cultural theories of medicine, estimates of complaint severity and likelihood of future pathology. While divergence from the original notion of a ‘sick role’ can be observed, parallels can be identified with Bury’s (1982) idea of illness (or in our case ‘worry’) as a ‘biographical disruption’ occasioning the need for medical intervention and narrative reconstruction (Williams, 1984) through advice. Most interestingly, the examination

of empirical data such as we have here suggests the social meaning of the problem is judged against internalised and largely implicit cultural norms—the ‘am I normal’ question—has assumed considerable significance in contributors’ accounts.

Discussion

The present paper has illustrated the role that studies informed by corpus linguistics can have for understanding subjective, culturally embedded and linguistically displayed notions of illness and health. In the study of health care encounters, whether over the internet or face to face, it offers researchers, practitioners and policymakers the opportunity to contextualise their insights within a much larger data set. Moreover, based on the experience of language learning, there is a growing role for corpora of data such as the present set in the study of health communication and in education for health professionals (Brown et al., 2006). This forms part of a broader movement which has been termed the ‘shift towards empirical methods’ (Sampson, 2005) in linguistics, where large naturally occurring data sets are increasingly preferred as a tool for the discovery of grammars, meanings and forms of usage. There may also be findings from corpus work that help us understand the relationship between language concerning health or bodily issues and how this has infused the language as a whole. Mishan (2004) reports some investigations of the occurrence of terms for body parts, averaged from the results of searches on two online corpora, namely the COBUILD Bank of English website, <http://www.cobuild.collins.co.uk>, and the British National Corpus online: <http://sara.natCorporation.ac.uk>. Here, for example approximately 45% of the occurrences of the word ‘hand’ were idiomatic (‘on the one hand’, ‘give a big hand to’, ‘out of hand’, ‘to hand’, etc.). Work on idioms and idiomatic usage in health care can therefore usefully be based on the evidence of corpus research.

One important function of corpus research is to characterise the nature of a language and examine the ‘nature of the realisations’ of particular language functions (Conrad, 2002). Hence, we have characterised the way that vocabularies and concepts were deployed by the correspondents to this particular website. To anyone working on health with young people in the medium of English, there could be valuable lessons to be learned. Of course, the material here does not necessarily tell us about

how geographic differences affect the form and content of the health messages, how, for instance, American or Australasian teenagers would talk about the issues, nor does the data collection process allow the questions and problems to be traced to particular age, gender, class, ethnic or faith subgroups. This would have to wait upon the creation of corpora where researchers have proactively collected demographic and contextual information to accompany the language itself.

With a scope and reliability of analysis not otherwise possible (Biber & Conrad, 2004), corpus linguistics is able to provide a nuanced explication of communication dynamics or 'linguistic signatures' directly associated with a variety of health care interventions or inputs. Not only can these be used to provide insights into the texture of the interaction in question, they can further be used to educate professionals and patients, potentially leading to better information exchanges and clinical outcomes. For example, the reticence noted by many authors when clients (and some professionals) talk about sex (British Medical Association, 2005; Mason, 2005; Stewart, 2005; Weijts et al., 1993) is clearly not apparent in all encounters and is in the present corpus often meticulously and explicitly articulated. It is possible that the frank and detailed health information provided by the Teenage Health Freak website (in a series of illustrative questions and answers) may well encourage users to communicate their concerns candidly and therefore, in part, account for the unreserved nature of their messages. Accordingly, further research that compares the linguistic content of the exchanges presented on the website and the emails provided by its users might prove fruitful in ascertaining whether adolescents' requests for health advice and information are indeed influenced by online templates. However, what our research highlights is the role, and value, of web based fora as a means of eliciting the sexual health problems of a group who have often been reluctant to consult practitioners, peers and others for personal health advice and information (Suzuki & Calzo, 2004). As a medium of communication, electronic interaction is typically distant and asynchronous, allowing participants to interact in a less-constrained way than when conversing face-to-face. Email is thus an effective medium for self-disclosure, with users, as this study has revealed, displaying remarkably high levels of apparent candour in the information they are willing to disclose on line (Baron, 1998, 2000). Thus the use of

electronic messaging is an important part of the accessibility of internet health resources (particularly services tailored to the needs of young people).

In thinking about how we might conceptualise the pattern of speech and writing about sexual and reproductive health, it might almost seem that the distinctive pattern of language use discovered here constitutes a kind of 'genre' of writing in that context (Trinch, 2001). In the sense that a genre is a 'class' or 'kind' of text, we can go some way towards delineating its features. Literary theorist Mikhail Bakhtin's (1986, p. 60) view was that genres 'reflect the specific conditions and goals' of their sites of production and reception. These goals are manifested 'through [the genre's] thematic content, the selection of the lexical, phraseological, and grammatical resources of the language, and above all through compositional structure' (Bakhtin, 1986, p. 60). Thus, correspondents to the website display not only a certain kind of vocabulary, but also a particular kind of evaluative stance towards what may be wrong. Questions of whether a feature of their emotional or physical topography is 'normal', and whether there is potential deviance from implicit standards of comfort or presumed normativity, are a source of 'worry'. Thus the combination of corpus work with scrutiny of the actual manifestations of key terms *in situ*, as advocated by Orpin (2005), helps to elicit not only vocabulary but the normative structures (Caboni et al., 2005) and 'situated-logics-in-use' (Hawes, 1977).

Future research might enable us to correlate measures of language such as we have just described with a variety of clinical outcomes. The systematic study of large bodies of language will explicate the many and diverse forms of health care language and will yield greater insight into the meaning of health care interaction. The promise of corpus linguistics is that it will allow a detailed analysis of a variety of health care language styles and interactions, which can then be utilised in communication training programmes. This creative synthesis between health care and corpus linguistics has the potential to provide a wide variety of health practitioners with the information they need to make substantial improvements in care delivery in a range of settings. Corpus studies such as these have identified recurring patterns of medical interaction and, consequently, provide medical educators with linguistic information with which to inform communications training and development.

Acknowledgements

We would like to thank the two anonymous reviewers for their trenchant and constructive comments and ideas.

References

- Ackard, D., & Neumark-Sztainer, D. (2001). Health care information sources for adolescents: Age and gender differences on use, concerns, and needs. *Journal of Adolescent Health, 29*, 170–176.
- Adolphs, S., Brown, B., Carter, R., Crawford, P., & Sahota, O. (2004). Applied clinical linguistics: Corpus linguistics in health care settings. *Journal of Applied Linguistics, 1*, 9–28.
- Ang, I. (1985). *Watching Dallas: Soap opera and the melodramatic imagination*. London: Routledge.
- Aynsley-Green, A., Barker, M., Burr, S., Macfarlane, A., Morgan, J., Sibert, J., et al. (2000). Who is speaking for children and adolescents and for their health at the policy level? *British Medical Journal, 321*, 229–232.
- Bakhtin, M. M. (1986). *Speech genres and other late essays* (C. Emerson & M. Holquist, Eds. & trans.). Austin, TX: University of Texas Press.
- Barkham, P. (2000). Is the Net healthy for doctors? *The Guardian*, June. Available on the World Wide Web at: <http://www.guardian.co.uk/netprivacy/article/0,,329729,00.html>, Accessed on 10/08/06.
- Baron, N. (1998). Letters by phone or speech by other means: The linguistics of email. *Language and Communication, 18*, 133–170.
- Baron, N. (2000). *Alphabet to email: How written English evolved and where it's heading*. London and New York: Routledge.
- Bedford, E. (1957). Emotions. *Proceedings of the Aristotelian Society, 57*, 281–304.
- Biber, D., & Conrad, S. (2004). Corpus-based comparisons of registers. In C. Coffin, A. Hewings, & K. O'Halloran (Eds.), *Applying English grammar: Functional and corpus approaches* (pp. 40–56). London: Arnold.
- Biddle, L., Gunnell, D., Sharp, D., & Donovan, J. L. (2004). Factors influencing help seeking in mentally distressed young adults: A cross-sectional survey. *British Journal of General Practice, 54*, 243–253.
- Boekeloo, B. O., Schamus, L. A., Cheng, T. L., & Simmens, S. J. (1996). Adolescents' comfort with discussion about sexual problems with their physician. *Archives of Paediatric and Adolescent Medicine, 150*, 1146–1152.
- Borkovec, T. D., Ray, W. J., & Stfber, J. (1998). Worry: A cognitive phenomenon intimately linked to affective, physiological, and interpersonal behavioural processes. *Cognitive Therapy and Research, 2*, 561–576.
- Borzekowski, D., & Rickert, V. (2001). Adolescent cybersurfing for health information: A new resource that crosses barriers. *Archive of Paediatric and Adolescent Medicine, 155*, 813–817.
- Bostock, W. W. (2002). Atrocity, mundanity and mental state. *Journal of Mundane Behaviour, 3*(3), 351–365.
- British Medical Association. (2005) *Sexual Health, June 2005*, available online at <http://www.bma.org.uk/ap.nsf/Content/sexualhealthjune05>, Accessed 13/3/2006.
- Brosschota, J. F., Gerinb, T. W., & Thayerc, J. F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *Journal of Psychosomatic Research, 60*, 113–124.
- Brown, B., Crawford, P., & Carter, R. (2006). *Evidence based health communication*. Buckingham: Open University Press.
- Bury, M. (1982). Chronic illness as biographical disruption. *Sociology of Health and Illness, 4*, 167–182.
- Caboni, T. C., Braxton, J. M., Deusterhaus, M. B., Mundy, M. E., McClendon, S. A., & Lee, S. D. (2005). Toward an empirical delineation of a normative structure for college students. *The Journal of Higher Education, 76*, 519–544.
- Car, J., & Sheikh, A. (2004). Email consultations in health care: 1—Scope and effectiveness. *British Medical Journal, 329*, 435–438.
- Cohler, B. J., & Hammack, P. L. (2007). The psychological world of the gay teenager: Social change, narrative, and “normality”. *Journal of Youth and Adolescence, 36*, 47–59.
- Conrad, S. (2002). Corpus linguistic approaches for discourse analysis. *Annual Review of Applied Linguistics, 22*, 75–95.
- DeSocio, J. E. (2005). Accessing self-development through narrative approaches in child and adolescent psychotherapy. *Journal of Child and Adolescent Psychiatric Nursing, 18*(2), 53–61.
- Donzelot, J. (1979). *The policing of families*. New York: Pantheon.
- Drew, P. (1998). Complaints about transgressions and misconduct. *Research on Language and Social Interaction, 31*(3&4), 295–325.
- Erikson, E. (1959). *Identity and the life cycle*. New York: Norton.
- Foucault, M. (1990). *The history of sexuality*, Vol. 1. Harmondsworth: Penguin.
- French, D. P., Sutton, S. R., Marteau, T. M., & Kinmonth, A. L. (2004). The impact of personal and social comparison information about health risk. *British Journal of Health Psychology, 9*, 187–200.
- Gergen, K. J. (1991). *The saturated self*. New York: Basic Books.
- Gray, N. J., Klein, J. D., & Noyce, P. R. (2005). The internet: A window on adolescent health literacy. *Journal of Adolescent Health, 37*, 243–247.
- Gray, N. J., Klein, J. D., Noyce, P. R., Sesselberg, T. S., & Cantrill, J. A. (2005). Health information-seeking behaviour in adolescence: The place of the internet. *Social Science & Medicine, 60*, 1467–1478.
- Harre, R. (1986). *The social construction of emotions*. Oxford: Basil Blackwell.
- Hawes, L. (1977). Towards a hermeneutic phenomenology of communication. *Communication Quarterly, 25*, 30–41.
- Henk, A. M. J. (Ed.). (2002). Editorial: Cybermedicine and e-ethics. *Medicine, Health Care and Philosophy, 5*, 117–119.
- Holland, J., Ramazanoglu, C., Sharpe, S., & Thomson, R. (1996). *The male in the head: Young people, heterosexuality and power*. London: Tufnell Press.
- Jones, R., Finlay, N., Simpson, N., & Kreitman, T. (1997). How can adolescents' health needs and concerns best be met? *British Journal of General Practice, 47*, 631–634.
- Kappahn, C., Wilson, K., & Klein, J. (1999). Adolescent girls' and boys' preferences for provider gender and confidentiality in their health care. *Journal of Adolescent Health, 25*(2), 131–142.
- Keck, M. E., & Sikkink, K. (1998). *Activities beyond borders*. Ithaca and London: Cornell University Press.

- Klein, J., & Wilson, K. (2002). Delivering quality care: Adolescents' discussion of health risks with their providers. *Journal of Adolescent Health, 30*, 190–195.
- Laakso, V., Niemi, P. M., Grönroos, M., Aalto, S., & Karlsson, H. (2005). The worried young adult as a primary care patient. *Family Practice, 22*, 406–411.
- Lee, R. B. (2003). *The Dobe Ju/'Hoansi* (3rd ed.). London: Wadsworth/Thomson Learning.
- Link, J. (2004). From the 'power of the norm' to 'flexible normalism': Considerations after Foucault. *Cultural Critique, 57*, 14–32.
- Lou, C., Zhao, Q., Gao, E., & Shah, I. H. (2006). Can the internet be used effectively to provide sex education to young people in China? *Journal of Adolescent Health, 39*, 720–728.
- Macken-Horarik, M. (2003). Appraisal and the special instructiveness of narrative. *Text, 23*, 285–312.
- Mason, L. (2005). 'They haven't a clue!' A qualitative study of staff perceptions of 11–14 year-old female clinic attenders. *Primary Health Care Research and Development, 6*, 199–207.
- McCarthy, M., & Carter, R. (1994). *Language as discourse: Perspectives for language teaching*. London: Longman.
- Michaud, P., Suris, J., McPherson, A., & Macfarlane, A. (2004). Alice in cyberland: Use and abuse of health websites by young people. *Italian Journal of Pediatrics, 30*, 198–204.
- Milliken, J. (1999). The study of discourse in international relations: A critique of research and methods. *European Journal of International Relations, 5*, 225–254.
- Mishan, F. (2004). Authenticating corpora for language learning: A problem and its resolution. *ELT Journal, 58*, 219–227.
- Monson, R., Jackson, C., & Livingston, M. (1996). Having a future: Sexual decision-making in early adolescence. *Journal of Paediatric Nursing, 1*, 183.
- Orpin, D. (2005). Corpus linguistics and critical discourse analysis: Examining the ideology of sleaze. *International Journal of Corpus Linguistics, 10*, 37–61.
- Payne, R. A. (2001). Persuasion, frames and norm construction. *European Journal of International Relations, 7*, 37–61.
- Plummer, K. (1995). *Telling sexual stories: Power, change and social worlds*. New York: Routledge.
- Popay, J., Bennett, S., Thomas, C., Williams, G., Gattrell, A., & Bostock, L. (2003). Beyond 'beer, fags, egg and chips'? Exploring lay understandings of social inequalities in health. *Sociology of Health and Illness, 25*, 1–23.
- Precht, K. (2003). Stance moods in spoken English: Evidentiality and affect in British and American conversation. *Text, 23*, 239–257.
- Rice, R. E. (2006). Influences, usage, and outcomes of Internet health information searching: Multivariate results from the Pew surveys. *International Journal of Medical Informatics, 75*, 8–28.
- Rogers, A., & Mead, N. (2004). More than technology and access: Primary care patients' views on the use and non-use of health information in the Internet age. *Health and Social Care in the Community, 12*(2), 102–110.
- Samaritans. (2004). *Young People and Suicide*. Available on the World Wide Web at: <www.samaritans.org.uk/know/information/information sheets/youngpeople/SISYoungpeople.doc>, Accessed on 17/10/06.
- Sampson, G. (2005). Quantifying the shift towards empirical methods. *International Journal of Corpus Linguistics, 10*, 15–36.
- Scott, M. R. (1996). *WordSmith: Software tools for Windows*. Oxford: Oxford University Press.
- Skelton, J., & Hobbs, F. (1999). Concordancing: Use of language-based research in medical communication. *The Lancet, 353*, 108–111.
- Skelton, J., Wearn, A., & Hobbs, F. (2002). 'I' and 'we': A concordancing analysis of how doctors and patients use first person pronouns in primary care consultations. *Family Practice, 19*, 484–488.
- Stewart, M. (2005). 'I'm just going to wash you down': Sanitizing the vaginal examination. *Journal of Advanced Nursing, 51*, 587–594.
- Stubbs, M. (1997). Whorf's children: Critical comments on critical discourse analysis. In A. Wray, & A. Ryan (Eds.), *Evolving models of language* (pp. 100–116). Clevedon: Multilingual Matters.
- Subrahmanyam, K., Greenfield, P., & Smahel, D. (2006). Connecting developmental constructions to the internet: Identity presentation and sexual exploration in online teen chat rooms. *Developmental Psychology, 42*, 395–406.
- Suzuki, L., & Calzo, J. (2004). The search for peer advice in cyberspace: An examination of online teen bulletin boards about health and sexuality. *Applied Developmental Psychology, 25*, 685–698.
- Thomas, J., & Wilson, A. (1996). Methodologies for studying a corpus of doctor–patient interaction. In J. Thomas, & M. Short (Eds.), *Using corpora for language research* (pp. 92–109). London: Longman.
- Thurlow, C., Lengel, L., & Tomic, A. (2004). *Computer mediated communication: Social interaction and the internet*. London: Sage.
- Trinch, S. L. (2001). Managing euphemism and transcending taboos: Negotiating the meaning of sexual assault in Latinas' narratives of domestic violence. *Text, 21*, 567–610.
- Weijts, W., Houtkoop, H., & Mullen, P. (1993). Talking delicacy: Speaking about sexuality during gynaecological consultations. *Sociology of Health and Illness, 15*, 295–314.
- Williams, G. (1984). The genesis of chronic illness: Narrative reconstruction. *Sociology of Health and Illness, 6*, 175–200.