"Can I Be at Risk of Getting AIDS?"
A Linguistic Analysis of Two Internet Columns on Sexual Health

Kevin Harvey (Nottingham), Miriam A. Locher (Basel), and Louise Mullany (Nottingham)

Abstract
Recent global statistics highlight that, out of all new cases of HIV infection, 45% are diagnosed in young people (UNAIDS 2008). Despite a range of new initiatives aimed at increasing young people's knowledge of HIV at the beginning of the first decade of the twenty-first century (UNAIDS 2001), latest figures highlight that such initiatives have not been wholly successful in preventing new infection in young people (UNAIDS 2011). In light of this, the language patterns that young people use when seeking information about HIV/AIDS are investigated. Our focus in particular is on computer-mediated-communication, a relatively under-researched area in the sphere of health communication. Building on previous research (Locher 2006, 2010; Harvey et al. 2008; Harvey 2013), we examine one UK and one US Internet-based, professional, health advice column as sources of advice-information for young people. Despite numerous established health campaigns, young advice-seekers' questions reflect misinformed conceptions, such as the conflation of HIV and AIDS and confusion as to the way in which the virus can be contracted. Our linguistic research gives access to young people's lay beliefs about sexual health and highlights the need to redress such beliefs, with the aim of improving the effectiveness of health education initiatives. We suggest that computer-mediated communication can be one effective medium through which to assess young people's knowledge about HIV/AIDS, as well as effectively disseminating sexual health advice and information by health care bodies.

1 Introduction
In 2001, the General Assembly of the United Nations agreed upon a Declaration of Commitment on HIV/AIDS which set the goal of ensuring that 95% of young people globally (where "young people" is defined as individuals aged between 15–24), would have comprehensive knowledge of HIV. However, latest statistics show that this target has not yet been achieved. One UNAIDS report illustrates that only 37% of young males and 28% of young females globally have even a basic knowledge of HIV (UNAIDS 2008), well below the commitment set in 2001. There is some good news. In 2010, cases of new infections had fallen to 15% less than in 2001 (UNAIDS 2011). However, in June 2011, statistics show that 15–24 year olds, and young women in particular are most at risk from HIV infection.

* We would like to thank the editor of Linguistik Online and the anonymous reviewers for their critical and helpful feedback that resulted in a much clearer text.
accounting for 22% of all new cases (UNAIDS 2011). These figures illustrate the importance of examining young people's perspectives on sexual health, in particular their attitudes towards and beliefs about HIV/AIDS, and the language through which they encode them. Evidence demonstrates that young people have many questions and concerns about their health (Jones et al. 1997) which they wish to discuss with professionals (Klein/Wilson 2002), but are reluctant to do this in a face-to-face context (McPherson 2005; Locher 2006; Harvey et al. 2008; Harvey 2013).

The topic of HIV/AIDS has provided the focus for much-needed linguistic enquiry in recent years. For example, Finn and Sarangi's (2009) work on the stigmas surrounding the illness in India, including a focus on the role of NGO talk and health promotional discourses (Finn/Sarangi 2008, 2010). Jones (2009) examines formulations of identity and how the discursive construction of expertise is manifested in online gay Internet forums. Additionally, Higgins and Norton (2009) emphasise the importance of examining how lay language interacts with authorised discourses about HIV/AIDS. Their edited collection deals with linguistic approaches to HIV/AIDS in a range of different geographical locations and provides insight into discursive formulations of attitudes towards HIV/AIDS and how this intersects with gender, poverty and institutional authority in a range of nation states.

In this corpus-based study quantitative and qualitative methods are combined to examine some of the linguistic choices made by contributors to two health websites when communicating problems about HIV and AIDS to professionals. The first website, entitled the Teenage Health Freak (THF), targets young people and offers a health forum tailored to the needs of those aged between 10–17. The second website, entitled Lucy Answers (LA), is a professional, US-based health advice column that targets slightly older individuals of university age (18 to mid twenties). The quantitative approach draws upon the tools and techniques of corpus linguistics. Corpus-based linguistic studies involve analysing extensive data sets, consisting of thousands or millions of words so as to discover patterns of use which may be then subjected to more interpretive analyses (Baker 2006). The approach involves interrogating large data sets or "corpora" of language and applies both quantitative and qualitative techniques to linguistic analysis; it is thus able to overcome some of the potential drawbacks of these approaches used in isolation. Corpus linguistics is an approach that is becoming increasingly popular in analysing health communication (Adolphs et al. 2004). The findings of the corpus study are then integrated with a more detailed qualitative analysis of specific data stretches where lexico-grammatical patterning is examined.

The next section of the paper focuses in more detail on computer-mediated communication and our selected websites in particular, along with a consideration of the importance and complexities involved in advice giving in health contexts. The mixed-methodological approach is then detailed in section 3. Results of the empirical data analysis are presented in section 4, and the consequences of these are discussed in section 5. A conclusion then follows.

2 Background: Computer-mediated communication and advice-giving

Both Teenage Health Freak and Lucy Answers are run by health professionals who provide credible, confidential and evidence-based advice and information. However, while the task of
the health professionals working on Lucy Answers is to target college students and the Teenage Health Freak website caters for adolescents specifically between the ages of 10–17, it is important to stress from the outset that there is no verifiable information of who actually reads and contributes to the sites – all contributors are anonymous. Lucy Answers does not ask for the disclosure of any personal identity information and has no means of following up information. The Teenage Health Freak contributors are given the option of self-disclosing age and sex, though there is also the option of choosing not to do this. Ninety percent of respondents choose to provide this information and they self-identify as young people aged between 10–17. Requests peak at those self-identifying as 13 and 14 years old. The mention of any names or other information that could potentially be used to identify individuals is removed by the website owners. This lack of identifying information is deemed to be one of the advantages of the online format. Research has shown that, as young people have their identities completely protected, they are more likely to ask questions that they may be unwilling to ask in a face-to-face settings with members of the medical profession, youth health workers or parents. Computer-mediated-communication thus provides an outlet for them to ask questions with their identities completely protected (see Locher 2006; McPherson 2005). The lack of identity information means that, as researchers, we cannot know with certainty that the age of the informants falls within the target age of the websites. However, for the purpose of our analysis, we work with the assumption that we are indeed dealing with young people, as do the medical practitioners that work on both websites. We will return to this point at a later stage.

There is much evidence of the awareness of HIV/AIDS among young people (Wight 1993a; Rosenthal/Moore 1994). Yet awareness does not necessarily constitute knowledge and understanding. Comparatively little is known about young people's knowledge and representation of HIV/AIDS and how, in particular, it is manifested and encoded linguistically through the language choices they make when seeking advice. This lack of linguistic detail represents a research gap. AIDS did not arrive with its own vocabulary and register: there have been linguistic difficulties right from the start (Koestenbaum 1990). Consequently, talking about HIV/AIDS is fraught with difficulties for any age group, with AIDS being described as a "linguistic battlefield" (Callen 1990: 171). For example, as Koestenbaum (1990: 164) suggests, AIDS is no mere acronym, it is rather "an accusation", something "more than simply a sickness". Moreover, the complexity of HIV/AIDS-related terminology exercises even health experts themselves, for whom the register can still constitute a source of ambiguity (Rafiquzzman 1995). Indeed, as Crystal (2001: 120) observes, the unabbreviated form of AIDS (Acquired Immune Deficiency Syndrome) is "so specialized that it is unknown to most people". Given the potential for the linguistic misrepresentation of HIV/AIDS, it is not surprising that standard-setting agencies such as UNESCO (2006) provide strictures concerning the use of terminology in HIV/AIDS communication, warning how the imprecise use of language is liable to reproduce myths about HIV/AIDS and engender unnecessary scaremongering.

In light of these linguistic tensions surrounding HIV/AIDS it is important to analyse the linguistic representation of HIV/AIDS, particularly because paying close attention to language can reveal much about people's perceptions about HIV/AIDS. As Allen (2007: 172) puts it, language is a powerful vehicle by which people communicate their values and beliefs.
Thus, by examining the linguistic choices through which young people in particular encode their concerns when seeking advice and information about sexual health, this study aims to explore their perceptions of HIV/AIDS representation, paying particular attention to concerns about transmission and contraction. These two processes have been most commonly associated with misunderstandings and knowledge shortfalls (Warwick/Aggleton/Homans 1988; Sikand/Fisher/Friedman 1996; Helman 2007).

In order to contextualize this research, the importance of computer-mediated communication and health discourse, as well as advice giving, are highlighted from a linguistic point of view. Richardson (2005: 1) reports that, in our "medialized world…[o]nce Americans have internet access, it turns out that finding health information is one of the most common ways in which they use it". Professionals working within health sectors have also clearly recognized the potential of online platforms to reach a wide number of their target audience, as evidenced in the large number of sites to be found on the web. There are sites run by public organizations such as the US National Institutes of Health (http://www.nih.gov/, accessed April 19, 2013) or the UK National Health Service (http://www.nhs.uk, accessed April 19, 2013), societies dedicated to information about illnesses (e.g., http://www.cancer.gov/, http://www.breastcancercare.org.uk/, both accessed April 19, 2013) but also advice sites targeted at particular groups of people (http://www.healthtalkonline.co.uk, http://www.youthhealthtalkonline.co.uk, both accessed April 19, 2013; see also Locher 2010, 2013). It is the latter type that is focused on in this paper.

Credible internet resources, carefully managed and maintained by reputable health care professionals, offer the following opportunities to health educators: they can reach a large number of people who are looking for advice; they can provide information at a mouse-click; this information can easily and quickly be updated; the target group can be given the possibility of searching archives (Locher 2010). Finally, as we have already highlighted above, in cases where questioners remain anonymous, they can look for information on delicate topics without fear of embarrassment (cf. Alexander 2003; Van Roosmalen 2000). Both sites under investigation in this paper are similar in making use of all of these advantages.

To seek and give advice or to impart information in an advice-column is not as straightforward and easy as it may initially sound (Locher/Limberg 2012). According to Searle (1969) and the Oxford English Dictionary, giving advice is more a "counsel" or "opinion" rather than an "order". In addition, advice-giving is a delicate act to carry out since there are "inescapable messages of authority, expertise and intimacy in advice" (DeCapua/Huber 1995: 128). According to Goldsmith and MacGeorge (2000), we can say that, in Anglo-Western contexts, giving and seeking advice is thus generally considered to be face-threatening (cf. the notion of "face" in Goffman 1967 and "face-threatening act" in Brown/Levinson 1987). If we think of these comments in light of the fact that the way in which we formulate our messages has an impact on the way messages are received, it becomes clear that health educators who use the form of online advice columns need to reflect about the relational aspect of the language used. This is the case because health educators have a clear mission to pursue and will thus strive for an adequate rendition of advice. Locher

---

1 This statement is based on the pewinternet.org reports (Madden 2003).
(2006, 2010, 2013), for example, showed how the ideal of non-directiveness influenced the linguistic way in which advice was given in *Lucy Answers*. Overall, there was a great deal of mitigation, realized by syntactic choices (questions and statements rather than imperatives), lexical hedges (*maybe, perhaps*), and giving options.

The people responsible for our advice columns thus engage in a delicate act when imparting information and advice to the target audience. While it is possible to study how advice is rendered in such sites in order to improve linguistic knowledge of this practice (cf., e.g., Locher/Hoffmann 2006; Locher 2006, 2010, 2013; Harvey et al. 2008), a focus can additionally be given to the analysis of the advice-seekers and the concerns they raise.

### 3 Methodology and data

The majority of previous research into young people's knowledge about and attitudes towards sexual health principally uses self-report methodologies (for example, Rosenthal/Mitchell 1996), which typically involve the use of questionnaires in order to explore the subject. Self-report respondents may decline to provide answers or substitute random replies for earnest responses (cf. Moore/Rosenthal/Mitchell 1996: 186). Additionally, respondents may display euphemistic constraint, under-representing their sexual behaviours and attitudes (cf. ibid.). In contrast, research is emerging where naturally-occurring, computer-mediated linguistic data provides the analytical focus (cf. Locher 2006; Harvey et al. 2008; Harvey 2013). In our case, all electronic data are authentic examples written by online contributors in non-experimental, non-researcher elicited settings before the research project began. Thus, rather than arising as a result of pre-determined researcher-set questions, the linguistic data used in this study have not been elicited in a context constructed/manipulated by the researcher. By distancing this study from self-report methodologies, this is not to suggest that the health messages to be analysed in this paper are entirely free of fabrication, understatement or overstatement. However, the fact that the question data are non-elicited for research in the first place means that the concerns communicate are principally motivated by what the contributors deem to be personally relevant, a factor which arguably helps to account for the often frank and non-euphemistic detail of their online questions and requests for help. Given the electronic contexts in which these contributors are free to submit their concerns online to the advice forums, without revealing their individual identities, the health messages analysed in this paper constitute a unique vantage point from which to survey young people's representations of contemporary sexual health, arguably giving more penetrating insights than are possible through more traditional, artificially-constructed methodological approaches such as questionnaires and/or structured interviews.

Our study combines both quantitative and qualitative approaches, taking a mixed-methodological approach to the study of naturally-occurring linguistic data. There has been an observable move towards adopting mixed-methodological approaches within linguistic study in recent years (Dörnyei 2007), and this mirrors a more general move towards a mixed-methodological paradigm within the social sciences more broadly (Whitley 2007).

From a quantitative perspective, we use a corpus linguistic approach to identify and describe language patterns that young adults use when formulating concerns about HIV/AIDS. We do this by conducting searches on particular lexical items as part of a register/vocabulary
analysis. Our analysis, in particular, focuses upon the verbs used to encode HIV/AIDS transmission and contraction.

This broad, quantitative survey is then enriched by a qualitative examination of the language used in the full surrounding context of the message. Lexico-grammatical patterning (Biber et al. 1999) is examined qualitatively through a closer study of the language used within a full stretch of discourse taken from the advice-seekers' messages. This combination of methodologies enables relevant exchanges to be identified so that these can then be analysed in depth. The linguistic insights which we gather from this combined methodological approach hopefully offer an alternative perspective to much research in young people's sexual health.

The data that constitutes this study will now be described in full detail. Lucy Answers (LA) is an online advice column based in the U.S. run by a professional team of health educators at a large American university. (The name of the site has been changed at the request of the site owners.) It has been in use since 1993 and its declared mission is as follows:

> to provide quality healthcare, 'by providing factual, in-depth, straight-forward, and non-judgmental information to assist readers' decision-making about their physical, sexual, emotional, and spiritual health'

(LA 2004)

The topics covered are relationships, emotional health, sexuality, sexual health, drugs, general health, fitness and nutrition.

Advice-seeking and advice-giving is organized in the form of published problem and response electronic 'letters' and thus can be associated with the genre of the classical newspaper advice column. The exchanges are stored in archives that are organized according to topic and that can be searched by the readers. In fact, the readers are encouraged to use the archive before submitting questions since only five questions a week are answered (there is an intake of 2000 questions a week). The site thus does not offer any emergency help. It is targeted at college students from the home university, i.e. at young adults, but is accessible to other people as well due to its public Internet location. As has been highlighted above in section 2, as the advice-seekers are entirely anonymized, there is no demographic information available on who actually constitutes the body of people. However, since only five of the hundreds of authentic questions that the team of health educators receives every week are published, it can be assumed that the team picks a selection that they consider to represent concerns that are relevant to the college target group. The advisors are a team of health educators, who publish under the female pseudonym Lucy.

The corpus available consists of the 2,286 question-answer pairs that can be accessed in the archive and that were published between 1993 and 2004 (Locher 2006). For this study, two sub-corpora are relevant. The first consists of all questions contained in the archive on "sexually transmitted diseases" (STDs; N=150, 6.5% of the overall corpus), which constitutes a section within the topic on 'sexual health' (N=469, 21% of overall corpus). The second consists of the 58 questions and responses within this sub-corpus on STDs that are relevant to the college target group.2 The potential drawback of this selection by the institution Lucy Answers is that we get to see questions that the health educators think are relevant to the target audience. In the case of Teenage Health Freak, we get all submitted questions.

2 The potential drawback of this selection by the institution Lucy Answers is that we get to see questions that the health educators think are relevant to the target audience. In the case of Teenage Health Freak, we get all submitted questions.
specifically indicated as dealing with "AIDS/HIV". This further selection corresponds to 2.5% of the overall corpus and to 38.7% within the STD corpus.

*Teenage Health Freak* is a British-based website run by a small team of medical doctors who specialise in child and adolescent health, reporting under the collective persona *Dr Ann*. It is named after one of the two founding doctors of the website, Dr Ann McPherson. The name is used as a collective female persona under which all members of the team publish their advice (see Harvey 2013; Mullany 2013). The site has been online since 2001. Its mission is articulated as follows:

> [t]o provide web-based, accurate and reliable health information to teenagers in a contemporary, cringe-free, entertaining and informative way. […] To produce relevant, electronic health information in a format that enables young people to take effective responsibility for their health related actions within an appropriate moral framework. To provide a database that can also be used by parents, teachers and health professionals.

(http://www.teenagehealthfreak.org 2001)

The target readership are adolescents aged 10–17, and the topics covered are equivalent to the ones in *Lucy Answers* – key topics are listed as sexual health, relationships, drugs, alcohol and smoking, the body, feelings, illnesses and weight and eating. As with *Lucy Answers*, young people are first instructed to look in the archive and then if they cannot find the information they require, they are encouraged to submit a question to *Dr Ann*. They type the advice-seeking request into a simple text box which is automatically submitted to the website. The site's interface is specifically designed to appeal specifically to a young readership, combining cartoon-style images with informal text. The website's text box section where young people are given the opportunity to ask a question is posed as follows: "Do you ever have questions that you wouldn't ever dare ask your parents? Then ask Dr Ann!" The website producers attempts to appeal to their targeted demographic can be clearly seen within the website design.

The contributors to the *Teenage Health Freak* website submit a question anonymously on the site in the hope to receive a publicly posted response. The site receives approximately 560 questions weekly, but similar to *Lucy Answers*, only a limited number are answered and published on the site. The data for this paper are taken from an overall corpus of all messages received on the site between 2004–2005, totalling a one-million word corpus of 62,000 messages. Within this corpus, 310 advice-seeking messages submitted by adolescents relating to the category "AIDS/HIV" were received by the website during this time. These 310 messages make up a 5,000 word sub-corpus for the *Teenage Health Freak* data that are focused upon in this paper and they are part of a larger set of questions on STDs (N=815). Similar to *Lucy Answers* (2.5%), the *Teenage Health Freak* HIV/AIDS messages make up a small amount of words in comparison to the corpus as a whole (0.5% of overall). The HIV/AIDS messages account for 38% of the overall number of messages posting on STDs, which again is remarkably similar to the percentage of overall STD messages in *Lucy Answers* (38.7%). The dataset contains a wide number of themes relating to HIV/AIDS and provides detailed insights into *Teenage Health Freak*'s contributors' beliefs about and attitudes to this important area of sexual health.

As illustrated above, while the corpora derived from the two sites are very similar, there are also clear differences. With *Lucy Answers* only the published question-answer pairs are
available, while in the case of Teenage Health Freak, the researchers obtained access to all of the questions submitted by advice-seekers (no matter whether they received published answers at a later stage). The two corpora also differ in the age of the target group, with Teenage Health Freak aiming at children and teenagers (10–17), and Lucy Answers aiming at college students. For the purpose of this study we assume that the majority of users of the sites correspond to the age group of the target readership, following the lead of the website owners who design their responses for these groups. There is also a difference with respect to the length of messages. On average, a question in Teenage Health Freak contains 16 words, while in Lucy Answers an average question is made up of 159 words. There are also potential socio-cultural differences in terms of where both websites are based, both with respect to the language used on the site and with respect to previous schooling that the contributors may have received on the subject of HIV/AIDS.

However, despite these caveats, both Lucy Answers and Teenage Health Freak are globally accessible web interfaces and monitoring exact usage in terms of geographical demographic is not possible with any degree of accuracy, particularly due to the websites' privacy policy. That said, although we cannot make precise demographic claims about the nature of the users, as an English language-based resource it is likely to be accessed by a wide range of English-speaking people, predominantly from Anglophone countries. In the case of Teenage Health Freak, although it is likely that the key users are based in the UK, there is also clear evidence of North American English lexis being used within some messages. Although there may be differences in sociocultural linguistic usage, as English language resources fulfilling the same medical advice-seeking and advice-giving functions, we believe our datasets to be comparable. Additionally, both sites are similar in their obvious educational aim and in the means chosen to impart this knowledge: an online advice column. They have, moreover, chosen a female advisor persona, and they both provide an archive function.  

In this study the following research questions will be explored:

1. What type of STDs are highlighted as concerns by the contributors within both corpora?
2. What concerns do the questioners raise in connection with HIV/AIDS?

4 Results

Our first research question, focusing on the exact type of STDs that are highlighted by the contributors within both corpora is motivated by our interest to examine how HIV/AIDS compares to other areas of sexual health concern. In order to study this, we looked at the corpus of STD messages for Teenage Health Freak (N=815) and the archive of "sexual health" questions in Lucy Answers (N=150). Table 1 presents the mention of specific sexually

---

3 Since we are dealing with two corpora of naturally-occurring data that are characterized by their own particular history, we cannot control the composition of the datasets to the same degree as corpus linguists usually do.
transmitted diseases (STDs) in both corpora, derived from a comparison of frequency word lists. HIV, AIDS and herpes range in the top for both corpora. If HIV and AIDS are combined, the category exceeds all other STDs in frequency in both corpora (Lucy Answers: 34.3%; Teenage Health Freak: 38.6%). Although the overall number of messages posted on HIV/AIDS are very small in comparison with both corpora as whole, it is thus clear that when questions about STDs are asked, HIV/AIDS figures as a prominent concern. In order to learn more about where knowledge gaps could be and to assess problems relating to conveying accurate information about contraction and transmission, these messages deserve further investigation.

<table>
<thead>
<tr>
<th>Word</th>
<th>Lucy Answers (N=150 questions)**</th>
<th>Teenage Health Freak (N=815 questions)***</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV</td>
<td>56</td>
<td>114</td>
</tr>
<tr>
<td>HERPES</td>
<td>55</td>
<td>96</td>
</tr>
<tr>
<td>WARTS</td>
<td>32</td>
<td>84</td>
</tr>
<tr>
<td>AIDS</td>
<td>27</td>
<td>209</td>
</tr>
<tr>
<td>YEAST infection</td>
<td>19</td>
<td>7.9</td>
</tr>
<tr>
<td>HSV</td>
<td>18</td>
<td>7.4</td>
</tr>
<tr>
<td>HPV</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>HEPATITIS</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>CHLAMYDIA</td>
<td>5</td>
<td>81</td>
</tr>
<tr>
<td>CRABS</td>
<td>3</td>
<td>103</td>
</tr>
<tr>
<td>GONORRHEA</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>KS*</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>PAPILLOMA</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>CIRRHOSIS</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>CONDYLOMAS</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>CONDYLOX</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>SYPHILLIS</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>THRUSH</td>
<td>1</td>
<td>106</td>
</tr>
</tbody>
</table>

* Kaposi's sarcoma;
** The same condition could be mentioned more than once within the same question and more than one condition could be mentioned within the same question.

Table 1: The occurrence of mention of STDs

Our second research question is "What concerns do the questioners raise in connection with HIV/AIDS?". Here we are interested in finding the similarities and differences between the two groups. Our close reading of the 58 Lucy Answers and 310 Teenage Health Freak questions on HIV/AIDS showed that five content types are used in the corpora:

1. Terminology and conceptual definitions of the terms
2. Concerns regarding transmission and causation
3. Symptoms and the likelihood of having HIV/AIDS
4. Psychological concerns
5. Testing technicalities

Illustrations of each question type are given in (1) to (5) in the same sequence and are taken from *Lucy Answers*:

(1) Dear Lucy, What is AIDS? -- Wondering

(2) Lucy -- I must sound paranoid, but all this fuzz about AIDs keeps me a little worried every time I go to get a haircut. Is it possible to get AIDS by the use of a contaminated razor? -- An AIDS paranoid

(3) Dear Lucy, Is it true that after contact with the AIDS virus, one develops flu-like symptoms within 6 weeks? -- Symptom monger

(4) Lucy, I am falling in love with a woman I know is HIV positive. We are great companions but there is also always an underlying sexual tension. I feel like I am playing with fire. What in the world should I do? -- Burning, but not consumed

(5) Dear Lucy, How long does it take to get results from an HIV test? -- Wondering

In order to quantify this finding, we assessed each question with respect to which type(s) of content concerns were raised. Table 2 lists the overall number of occurrences for each question type in both corpora. The results show that there is a difference in interest between the two age groups. According to these numbers, the questioners in *Lucy Answers* are not concerned with terminology and definitions or psychological queries. Instead, their main focus is on "transmission and causation" (74.6%), followed by the minor concerns of "symptoms and the likelihood of having HIV/AIDS" (10.2%) and "testing technicalities" (11.9%). The questioners in *Teenage Health Freak* have a wider range of concerns. Fifty percent of their questions are made up of either concerns regarding transmission and causation (22.9%) or concerns regarding terminology and definitions (24.8%). But they also ask about symptoms (19.1%) and psychological concerns (21%).

<table>
<thead>
<tr>
<th>Type of question</th>
<th><em>Lucy Answers</em> (N = 58 questions)</th>
<th><em>Teenage Health Freak</em> (N=310 questions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concerns regarding transmission and causation</td>
<td>44  74.6</td>
<td>72  22.9</td>
</tr>
<tr>
<td>Symptoms and the likelihood of having HIV/AIDS</td>
<td>6   10.2</td>
<td>60  19.1</td>
</tr>
<tr>
<td>Terminology and conceptual definitions of the terms</td>
<td>1   1.7</td>
<td>78  24.8</td>
</tr>
<tr>
<td>Psychological concerns</td>
<td>1   1.7</td>
<td>66  21.0</td>
</tr>
<tr>
<td>Testing technicalities</td>
<td>7   11.9</td>
<td>38  12.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>59  100.1</strong></td>
<td><strong>314  99.9</strong></td>
</tr>
</tbody>
</table>

Table 2: The type of questions concerning HIV/AIDS raised by the advice-seekers (double coding within a question is possible)

These differences might reflect greater understanding of (or a lack of willingness to question) fundamental terminological aspects of HIV/AIDS on the part of *Lucy Answers* contributors. The large share of questions on HIV/AIDS terminology and conceptual definitions on the part of the *Teenage Health Freak* contributors may reflect a lower level of sexual health literacy – or a willingness to question basic concepts (without fear of showing themselves up as being
Questions about HIV/AIDS terminology are justified, since even professionals have argued over (and still dispute) AIDS nomenclature. Such terms are not readily understandable. For example, there has been a drawn out and heated debate on the definition of both HIV and AIDS, with terms such as "HIV" passing through a range of contested acronymic incarnations before eventually settling into their present linguistic form. Yet even more recently confusion among professionals still subsists. Rafiquzzman (1995), for instance, has raised questions about the universal understanding of the AIDS acronym, highlighting its potential for ambiguity even among experts. It is not unsurprising, therefore, that such complex and loaded terminology is liable to cause problems among young lay advice-seekers and thus their terminology-based questions are perhaps to be expected.

Our next question, "How do the questioners refer to HIV/AIDS? What verbs do they use in combination with HIV/AIDS?" deals with how exactly the questioners refer to HIV/AIDS and, more specifically, what verbs they use in combination with HIV/AIDS. Once again we draw on the 58 questions on HIV/AIDS in Lucy Answers and the 310 questions on HIV/AIDS in Teenage Health Freak to explore this issue. To illustrate, there are three examples from Lucy Answers and three examples from Teenage Health Freak:

(6) Dear Lucy, How does a man get infected with HIV through heterosexual vaginal intercourse? Just wanna know (LA)

(7) Dear Lucy, I have been asked by my hairdresser about the risks of contracting HIV when using the facilities. What are the precautions one should take for this problem? Are there any scientific papers addressing this question? -- Miguel (LA)

(8) Hi Lucy, I am sure this question of mine may sound stupid and you have been asked a number of times. However, for me it is a very important question relating to my sexual life. My question is: does kissing, with sucking your partner's tongue and lips, transmit HIV? For me, sex without such kissing is no fun! Lately my girlfriend was told by someone that such kissing is risky, and therefore she refuses to give kisses during sex. Waiting anxiously for your reply. Signed- No kiss no fun

(9) I heard that drug users can get AIDS from using needles is this true (THF)

(10) Can you catch HIV if you wear an earring that might have been worn by somebody else before? (THF)

(11) Can you be born with HIV or AIDS or do you have to catch it? (THF)

Table 3 shows that both the questioners in Lucy Answers and Teenage Health Freak often mention HIV and AIDS but not in connection with a dynamic verb (that is, a verb that describes a process that commences and finishes). This is the "no mention of verb" category (as exemplified in examples (1), (4) and (5) above). The verbs absorb (virus), be infected/infection (no specification), catch, contract, get, get infected with, pass on, spread,
suffer from, and transmit co-occur with HIV/AIDS. Transmit, contract and get are the verbs most often used.

<table>
<thead>
<tr>
<th>Lucy Answers (N = 58)</th>
<th>Teenage Health Freak (N=310)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td><strong>%</strong></td>
</tr>
<tr>
<td>'no mention of a verb'</td>
<td>22</td>
</tr>
<tr>
<td>transmit HIV</td>
<td>10</td>
</tr>
<tr>
<td>contract HIV</td>
<td>7</td>
</tr>
<tr>
<td>get AIDS</td>
<td>4</td>
</tr>
<tr>
<td>be infected / infection (no specification)</td>
<td>4</td>
</tr>
<tr>
<td>get HIV</td>
<td>3</td>
</tr>
<tr>
<td>pass on HIV</td>
<td>2</td>
</tr>
<tr>
<td>spread AIDS</td>
<td>2</td>
</tr>
<tr>
<td>contract AIDS</td>
<td>1</td>
</tr>
<tr>
<td>absorb (virus)</td>
<td>1</td>
</tr>
<tr>
<td>catch (a disease)</td>
<td>1</td>
</tr>
<tr>
<td>get infected with HIV</td>
<td>1</td>
</tr>
<tr>
<td>transmit AIDS</td>
<td>1</td>
</tr>
<tr>
<td>suffer from AIDS</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>60</strong></td>
</tr>
</tbody>
</table>

Table 3: Verbs used in combination with HIV/AIDS by questioners (more than one occurrence within a question is possible)

The questioners in Lucy Answers clearly keep HIV and AIDS apart. Transmit HIV and contract HIV are the most commonly used collocations with this group. In contrast to Lucy Answers, the questioners in Teenage Health Freak might be liable to conflate HIV and AIDS. This is not, of course, to suggest that the adolescents do not know the difference between HIV and AIDS, but that AIDS is sometimes spoken about as being an immediately acquired and fully-developed infection, rather than a syndrome that develops over time, for example "can I catch AIDS off my cat?" Collapsing the distinction between HIV and AIDS in this way potentially results in confusion and reinforces "unrealistic and unfounded fears" (Watney 1989: 184) on the part of individuals who may well mistakenly believe themselves to be at risk of AIDS but not HIV. Such extreme worse case scenarios conceive of AIDS as something that sets in immediately after infection, a unitary phenomenon rather than a collection of different medical conditions. Such beliefs obscure, if not efface altogether, the existence of the virus HIV, which is indeed infectious. Such potential erroneous conflation of HIV infection with AIDS (by definition, the stage of HIV infection "when a person's immune system can no longer cope" (Terence Higgins Trust 2007: 1)) repeats some of the early and fundamental misconceptions about AIDS that were widespread during the 1980s and 1990s (Sikand/Fisher/Friedman 1996; Helman 2007), misconceptions which, as shall be shown shortly, are apparent in a further number of the teenagers' HIV/AIDS concerns.

The Teenage Health Freak contributors use get AIDS 52 times, and spread AIDS and contract AIDS in one instance each. To illustrate, consider examples (11) and (12):

(12) i want to know if AIDS can gotten through kissing? (THF)
The use of the verb *catch* appears thirteen times and requires special attention. *Catch* is an emotive word which potentially suggests a misconception about HIV transmission. Biber at al. (1999: 361) describe "catch" as an "activity verb", a verb denoting actions and events "that could be associated with choice". As the above examples illustrate, "catch" implies specific notions of agency on the part of subjects in the sense that it is within their power to avoid or prevent infection, with responsibility framed in terms of both general or universal agency (encoded via the second-person: "Can you catch...") individual control via the first-person singular pronoun: "Can I catch...". As Johnson and Murray (1985: 152) put it, "catching" an ailment (as in catching a cold) semantically implies a degree of co-operation: "We catch things...in ways which are our own fault; we blame ourselves – we should have worn galoshes, and should not have sat in a draught".

However, despite the "measure of participation" (Fleischman 1999: 10) entailed by use of the verb "catch" (and with this the implicit acknowledgement of responsibility on the part of the contributors for maintaining their sexual health), many commentators and public health bodies stress that neither HIV nor AIDS can be "caught" (Watney 1989: 184). In the health messages in which *catch* appears, for example, the various and recurring realisations of the word unavoidably and infelicitously conjure notions of the common cold and influenza, as evidence from the British National Corpus (BNC) attests. Consulting the 100 million word British National Corpus, a corpus representative of both spoken and written English language as a whole, reveals that, as a transitive verb, *catch* co-occurs with the direct objects *cold* (113), *chill* (21), *bug* (20) and *colds* (7). As this range of terms indicates, one typical use of the verb *catch* in general English is to describe the acquiring of relatively minor and common infections – minor in the sense of their being widespread and generally innocuous (though *bug*, of course, potentially relates to more serious infections such as MRSA, the so-called *super bug* (Knifton 2005)). With regard to more serious viruses and illnesses, other less euphemistic constructions are used in the BNC to describe the process of becoming infected and the onset of morbidity: for example, HIV is typically *contracted*, *got*, *acquired*, while AIDS is *got*, *developed*, *contracted*. In general English, therefore, the use of *catch* to signify becoming infected is routinely associated with common and relatively trivial infections and illnesses.

Given this association, a corollary of using *catch* to describe infection with HIV/AIDS is, unavoidably, to encode the assumption that the virus can be acquired via casual contact, possessing a transmission efficacy similar in effect to both colds and influenza. As such the adolescents' talk of "catching" HIV/AIDS figuratively transforms the virus from something which is, in reality, difficult to transmit and is only communicable via specific routes (Patton 1986: 142; Terence Higgins Trust 2007: 2) to something highly contagious, liable to spread rapidly and extensively.

Our corpus evidence suggests that some of the *Teenage Health Freak* contributors may well be liable to employ more inappropriate and inaccurate language forms when describing the process of transmission (although it has to be recognised that some of differences in the way
the two sets of users speak about HIV/AIDS could also be due to the general way in which these two culturally distinct groups commonly discuss health issues). This information is potentially important for health educators to keep in mind. Although we cannot, categorically, explain why the Teenage Health Freak contributors appear to be working with such misconceptions, we suggest that this age group, as opposed to the older young adults using Lucy Answers, may be subjected more to school rumour and folkways surrounding sexual health and may have generally less experience in being sexually active. The finding may also reflect the inefficiency of school education initiatives tailored toward teenagers – courses which do not take into account the misconceptions that adolescents work with and through which they are liable to filter official 'correct' information (Aggleton/Homans 1987; Wight 1993b; Helman 2007). For this reason we believe it is important to point out such misconceptions. Having proposed these interpretations, however, we need to recall that we are without precise demographic information on the two groups of contributors and can therefore only hypothesize on whether the age and education factors led to the clearly different results we can report here.

The final research question, "How do the respondents refer to HIV/AIDS? What verbs do they use in combination with HIV/AIDS?", acts as a point of comparison. We turn to how the respondents, i.e. the team behind Lucy and Dr Ann, refer to HIV/AIDS and in particular what verbs they use in combination with HIV/AIDS. Drawing on the responses to the 58 Lucy Answers questions that deal with HIV/AIDS and looking at the 14 published responses in the Teenage Health Freak archive on HIV/AIDS, Table 4 shows that there is a clear preference in Lucy Answers for transmit HIV (not AIDS) / the passive form infected with HIV (not AIDS) / and contract HIV (not AIDS), while Dr Ann is less consistent in this usage.4

<table>
<thead>
<tr>
<th></th>
<th>Lucy Answers (N = 58)</th>
<th></th>
<th>Teenage Health Freak (n=14)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmit HIV</td>
<td>58</td>
<td>52.3</td>
<td>6</td>
<td>42.8</td>
</tr>
<tr>
<td>'no mention of verb'</td>
<td>17</td>
<td>15.3</td>
<td>2</td>
<td>14.2</td>
</tr>
<tr>
<td>Infection</td>
<td>14</td>
<td>12.6</td>
<td>1</td>
<td>7.1</td>
</tr>
<tr>
<td>contract HIV</td>
<td>11</td>
<td>9.9</td>
<td>2</td>
<td>14.2</td>
</tr>
<tr>
<td>catch HIV</td>
<td>3</td>
<td>2.7</td>
<td>2</td>
<td>14.2</td>
</tr>
<tr>
<td>pass HIV</td>
<td>1</td>
<td>0.9</td>
<td>2</td>
<td>14.2</td>
</tr>
<tr>
<td>develop AIDS</td>
<td>1</td>
<td>0.9</td>
<td>2</td>
<td>14.2</td>
</tr>
<tr>
<td>get HIV</td>
<td>1</td>
<td>0.9</td>
<td>2</td>
<td>14.2</td>
</tr>
<tr>
<td>get aids</td>
<td>1</td>
<td>0.9</td>
<td>2</td>
<td>14.2</td>
</tr>
<tr>
<td>affected with HIV/AIDS</td>
<td>1</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cause AIDS</td>
<td>1</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>get HIV infection</td>
<td>1</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>get HIV/AIDS</td>
<td>1</td>
<td>0.9</td>
<td>1</td>
<td>7.1</td>
</tr>
<tr>
<td>receive HIV</td>
<td>1</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>1</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>111</td>
<td>100</td>
<td>16</td>
<td>99.6</td>
</tr>
</tbody>
</table>

Table 4: Verbs used in combination with HIV/AIDS by respondents

In addition (and beyond the analysis of the two sub-corpora just mentioned), we noticed that Lucy and Dr Ann mention HIV and AIDS in many more responses on STDs and other

---

4 Though a larger sample of Dr Ann responses would be needed to assess this more thoroughly.
questions than was requested by the questioners. In other words, the sites take the opportunity to pass on knowledge about HIV/AIDS also in responses to questions that did not explicitly enquire about HIV/AIDS. This is the case in (13), where the question is about the use of hard drugs:

(14) Lucy -- Recently my boyfriend began injecting cocaine. I've noticed that he and his friends share their needles, but 'clean' them first with bleach and water. Is this a valid way to avoid contamination? Signed -- Worried and Wondering

Dear Worried and Wondering, Oh, boy! There's quite a few levels to this question. First of all, cleaning shared needles can be an effective way to prevent transmission of HIV (Lucy assumes that this is what you mean by contamination). But, by putting cleaning in quotation marks, you put into question its definition. To prevent HIV transmission, your boyfriend and his friends must clean the complete works, including syringe and cotton, not solely the needle. Bleach should be drawn into the syringe, and then emptied and flushed with water, at least three times. A recent news report questioned the effectiveness of this technique. The NYC AIDS hotline still recommends it, although there is no scientific research to support either conclusion. Lucy would like to ask you to think about some things... How does your boyfriend's drug use affect you -- emotionally? physically? Do you use a condom every time you have sex? Is he aware of the potential dangers of shooting coke -- Are you? It's different than snorting it, or free basing. Lucy will stop here because you didn't ask her about these things. But, if you need a place to talk, feel free to call Counseling and Psychological Services (CPS) at <number>, and make an appointment with a counselor.

This strategy can be called 'widening the scope of the answer' and can be explained the public dimension of the site which is geared towards reaching a wider readership with relevant information (cf. Locher 2006, 2010; Locher/Hoffmann 2006). This widening also occurs in responses to HIV/AIDS questions. Lucy Answers and Dr Ann take the opportunity to mention and draw attention to other matters which are related to the initial concern. The following example from Teenage Health Freak further illustrates the phenomenon:

(15) how do you get rid of aids

Dear "how can you get rid of aids", Aids, which stands for Acquired Immune Deficiencey Syndrome, is caused by the virus called the Human Immunodeficinecy Virus and once a person is infected it is for life. At present there is no cure for AIDS but there are drugs which help to keep it under control so it is very important to do everything possible not to get the virus in the first place. Talk to your partner before you decide to have sex to make sure you are both ready for sex and understand about AIDS and other sexually transmitted infections. Always always use a condom when you have sex which will help protect you from getting the virus if your partner is HIV positive. If you have unprotected sex or have injected drugs using a needle that someone else has used who has the virus make
sure you get checked out by having a free blood test available from your family doctor or a clinic.

Responding to the contributor's question directly concerning AIDS, Dr Ann further widens the scope of the answer by drawing attention to other STDs besides HIV/AIDS, thereby further reinforcing the importance of using a condom during sexual intercourse. In addition, Dr Ann takes the opportunity to provide advice more broadly related to the consequences of taking part in sexual relationships. In widening the scope of the response here, the health advisor supplies additional information about HIV/AIDS (the importance, for instance, of taking a blood test) that was not directly asked for by the teenager in the first instance, advice which is nonetheless still relevant to the contributor's question.

5 Discussion

Our corpus-driven linguistic study has identified the kinds of information that contributors of two websites aimed at young people seek and has described the lexical patterns used when the contributors are seeking information about HIV/AIDS, in particular patterns and commonalities surrounding the transmission and acquisition of HIV/AIDS. Our findings suggest that, despite the presence of established health campaigns tailored to young people, the advice-seekers' questions often reflected misinformed conceptions and folk beliefs regarding HIV/AIDS, including confusion as to the precise mechanisms by which HIV can be transmitted and contracted. With the caveats highlighted above in mind, the Lucy Answers contributors appeared to be better informed on the nature of HIV and AIDS, their questions evincing fewer misconceptions and knowledge deficits. In contrast, the Teenage Health Freak cohort appeared to possess a range of misconceptions about HIV and AIDS, among which was the tendency for a number of the contributors to conflate the two terms or use them synonymously, potentially reinforcing the idea that HIV and AIDS are the same thing. As argued above, such a conflation might obscure awareness of the ways in which HIV is transmitted, thereby potentially impeding accurate assessment of degrees of risk in relation to young people's sexual behaviour (Aggleton et al. 1989: 59).

This paper has demonstrated that awareness of HIV/AIDS does not necessarily constitute knowledge and understanding. Thus the findings in this paper add impetus to the call for evidence-based sex education programmes that provide "full and factual information" (Independent Advisory Group on Sexual Health and HIV 2007: 12). The findings further highlight that, if health education initiatives are to change behaviours, they also need to address lay beliefs about HIV/AIDS, since it was clear from our linguistic analysis that the Teenage Health Freak contributors operated with folk conceptualisations of sexual health. Responding to lay beliefs like these is crucial because people are liable to filter official health education messages through popular beliefs about health (Herzlich/Pierret 1986; Wight 1993b; Helman 2007), reinterpreting them to suit their own needs (Aggleton/Homans 1987: 25). This is particularly the case with the subject of HIV/AIDS, where the pervasiveness of irrational myths that contradict scientific facts appeared to inform the work the adolescents did in order to try to make sense of this "complex, puzzling and quite terrifying phenomenon" (Treicher 1989: 34).
However, it is important to emphasise that the findings from our study cannot simply be generalised to the adolescent population and young people more widely. The people who contribute to the Teenage Health Freak and Lucy Answers websites constitute a specific group of anonymous users who articulate certain feelings and thoughts, and therefore their potential knowledge gaps, unique beliefs and communicative repertoires cannot be said to be representative of young people more widely. Mention of this caveat is also particularly important since we do not have access to demographic information about the contributors, which would confirm that we are indeed only dealing with young people or language users of a particular variety of English or educational background. In addition, many young people are well-informed about matters of sexual health and have no need to seek online health advice. Not all young people have questions about HIV and AIDS.

In drawing attention to the linguistic repertoires of the two sets of contributors in our study, the results from our corpus-driven analysis have potential practical relevance for health practitioners and educators concerned with the health of young people. The corpus approach affords an effective means of identifying the "incremental effect" (Baker 2006: 13) of linguistic patterns and commonalities in young people's understanding of sexual health. For example, the systematic uncovering of the recurring use of verbs such as get and catch affords an insight into attitudes towards and beliefs about HIV/AIDS transmission and infection. The subtle presence and full implication of these linguistic choices in the communicative routines of the contributors may well be overlooked by health professionals and policymakers unaccustomed to considering the significance of language in the shaping of interventions and policy. If educational initiatives are to be successful, then language, as Cameron and Kulick (2003: 154) argue, must not simply be regarded as "a medium for sex and health education but something that must be discussed explicitly as part of the process".

Furthermore, a linguistic approach offers a fresh methodological and analytical perspective in the domain of sexual health communication and represents a response to a call for different modes of research into sexual behaviour (Dockrell/Joffe 1992). Such a call underscores the difficulties associated with generating an adequate depiction of young people's sexual health, a challenge which owes, in no small part, to the fact that the right to privacy is, understandably, jealously guarded in this sensitive area of research (Moore/Rosenthal/Mitchell 1996: 186). It is not surprising, therefore, that, despite the increasing amount of surveys and epidemiological work on the subject, there is a lack of research concerning how people communicate sexual issues in naturally-occurring situations (Silverman 1997). Given the context in which young people are invited to submit their concerns to the online Teenage Health Freak and Lucy Answers websites, a naturally-occurring form of online advice-seeking, the electronic messages we have examined constitute a unique vantage point from which to survey contemporary youth sexual health, complementing more traditional methodological approaches such as questionnaires and interviews, often with problem-focused agendas dictated by researchers.

6 Conclusion

In conclusion, our comparative study has highlighted the value of the web-based forum as a means of eliciting the linguistic choices used to articulate concerns about HIV/AIDS from a generation who have often been reluctant to consult practitioners and others for sexual health
advice (Suzuki/Calzo 2004). The importance of computer-mediated health provision as an additional information source is evidenced in the vast number of health-related questions submitted to the two sites, including the contributors' HIV/AIDS related questions. This points to the importance of this type of electronic resource for young people. Research into sexual health and sexuality has described the communicative difficulties that participants (both patients and professionals alike) experience when discussing this highly sensitive topic face-to-face (Weijts et al. 1993; Reeves et al. 2004; Stewart 2005; Locher 2006; Pollock 2007; Emslie et al. 2007). Although not confined to the age group alone, the problem is particularly acute for young people, who consistently report poor communication between themselves and doctors (Resnick 1982; Malus et al. 1987; Jacobson et al. 1996; Beresford/Sloper 2003). The possibility to ask questions in an anonymous and easily accessible way may partly explain the success of the sites (Van Roosmalen 2000; Alexander 2003; Locher 2006).

Consequently, we believe that the educational and therapeutic advantages of creating online "personas" such as Lucy or Dr Ann merit further investigation, particularly as a means of reaching out to young individuals who might be reluctant to seek health advice directly from more traditional sources of professional support. In communicating their concerns electronically, the contributors to the websites in this study have taken the calculated step of seeking an alternative, non-interventional source of assistance. The online medium provides them with a venue in which they can deal with the same developmental issues as in their offline lives (Subrahmanyam/Greenfield/Tynes 2004). In the case of sexual health, the electronic forum provides "a safe place to vent feelings and express negative emotions" (Sullivan 2003: 86), a platform from which such delicate and exposing themes can be communicated in an anonymous and disembodied social context.

References


Jones, Roger et al. (1997): "How can adolescents' health needs and concerns best be met?". British Journal of General Practice 47: 631–634.


Lucy Answers: The address of the website cannot be disclosed at the wish of the institution.


