THE EVOLUTION OF ONLINE ASEXUAL DISCOURSE

BY

ANDREW CLINTON HINDERLITER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Linguistics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

       Associate Professor Marina Terkourafi, Co-Director, Chair
       Associate Professor Chilin Shih, Co-Director
       Associate Professor Corina Roxana Girju
       Associate Professor Randall W. Sadler

Abstract

Technological, social, and economic changes in recent decades have led to new possibilities for communication and for forming communities that are not tied to a specific geographical location. This creates new opportunities and challenges for studying language change, including the language of online communities. This dissertation provides such a case study by examining the development of, and changes in, online English language asexual discourse from the second half of the 1990s until late 2013, focusing on lexical items and multi-word expressions.

This dissertation combines three major research approaches—archival research, corpus research, and survey research . Using the Way Back Machine, databases of newspaper archives, and academic databases and references, historical conceptualizations of asexuality can be seen well before the emergence of asexual communities online, but I can find no evidence of asexual organizing prior to the 1990s. Largely using qualitative analysis of asexual websites, I give a historical account of the development of online asexual communities, and I argue that there have been at least two major conceptual shifts in the conceptualization of asexuality in the time period under consideration, which I call the "AVEN shift" and the "rise of intermediate categories."

I then discuss the construction of four corpora: I scrapped the largest asexual website (Asexual Visibility and Education Network [AVEN]) and a similar sized message board on a different topic as a control (*Non-asexual corpus*). I subdivided the AVEN data into two sub-corpora, based on the sub-forum topics. Most subforums are in the *AVEN-core corpus*. Some (e.g. "Just For Fun" or "Off-A") are mostly about topics other than asexuality, and were grouped as the *AVEN-other Corpus*. In addition, I scraped several asexual blogs and asexual communities other than AVEN (e.g. a LiveJournal community): these comprise the *Asexual-other corpus*.

Using a multinomial Naive Bayes classifier, I found moderate distinguishability between the AVEN-main and non-asexual corpura at the level of individual posts when only considering individual words. To rule out the possibility that the classifier was distinguishing AVEN vs. non-AVEN discourse, I first created an algorithm to remove from consideration words that probably refer to users. Second, I used the same classifier

on the AVEN-other corpora and the asexual-other corpora. Results for the asexual-other corpus are similar to results for the AVEN-main corpus, while results for the AVEN-other corpus are not. This suggests that the classifier is identifying asexual discourse vs. other discourse.

I used the AVEN-core Corpus to generate a list of "key-words" that well-characterize asexual discourse, and then investigate the evolution of three sets of these: intermediate category terms, romantic orientation terms, and the terms *repulsed* and *indifferent*. Results provide strong support for the "rise of intermediate categories" hypothesis, and also provide evidence of terminological change for the other two domains. To test the "AVEN shift" hypothesis, I conducted an online survey in Early 2012 about people's self-understanding prior to and after finding an online asexual community. Results provide evidence for all predictions of the "AVEN shift" hypothesis, although the changes were by no means monolithic.

Through this research, I illustrate the utility of applying methodologies from corpus linguistics and from machine learning for investigating the language of specific online communities. Further, I provide novel methodologies (or novel uses of existing methodologies) for problems likely to be faced by other researchers using corpus linguistics to study online language change in specific (sub)communities.

To my wife.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Technological, social, and economic changes in recent decades have led to new opportunities for communication and for forming communities that are not tied to a specific geographical location. This creates new opportunities and challenges for students of human language, including the study of language change, a topic that is of interest to both linguists and the general public alike.

Although the study of language change has long been a major part of linguistic theory, sound change has historically been preeminent, while other levels of linguistic analysis have received less attention. However, the same technological changes that have enabled online communication and communities—cheap, easily available computing speed, memory, and internet access—make it possible for linguists today to obtain, store, and process quantities of text far beyond what was possible until recently.

One of the research opportunities that has become possible is the study of specific online communities and how the discourse and terminology has developed over time in them. These can provide a sort of linguistic microcosm within which to study lexical evolution, where a huge amount of the relevant discourse is machine-readable, time-stamped, and in the public domain. This can also benefit the communities thus studied, as many members may be interested to learn more about the history of their group's discourse and terminology.

This dissertation provides such a case study by examining the development of, and changes in, online English language asexual discourse from the second half of the 1990s until late 2013, focusing on lexical items and multi-word expressions. In doing so, I illustrate the utility of applying methodologies from corpus linguistics and from machine learning for investigating

the language of specific online communities. Further, I provide novel methodologies (or novel uses of existing methodologies) for problems likely to be faced by other researchers studying online language change in specific (sub)communities.

This time period is chosen because there has been a growth of online asexual communities during it, and I have been unable to locate any pre-internet asexual organizations perceived as being, in some sense, continuous with online asexual communities. In these online communities, asexuals most commonly define asexuality in terms of not experiencing sexual attraction (Brotto et al., 2010; Scherrer, 2008), which reflects the definition of "asexual" given on the front page of the largest online asexual community, the Asexual Visibility and Education Network (AVEN): "An asexual is a person who does not experience sexual attraction." [1]

The English language online asexual communities I am aware of have a user-base primarily consisting of people in Anglo-American countries, especially the US, the UK, and Canada. Following general patterns of globalization, since the birth of English language online asexual communities, there has been a growth in online asexual communities in other languages (mostly Germanic and Romance languages), and these often cite (and sometimes translate) English language materials. Some of these are or were hosted on the AVEN server. Although worthy of linguistic investigation, these non-English language communities are beyond the scope of the present study.

There is nothing unique to asexual discourse about the fact that it has developed its own terminology. A great many professions have their own professional jargon consisting of group-specific lexical items and sometimes even syntactic constructions (e.g. lawyers' use of "whereas". Hock, 1991) and some subcultures on the margins of society develop secret languages (argots). Rather, what the study of online asexual discourse contributes to linguistic theory is two-fold. First, it provides a contemporary case study in the (ongoing) development of a discourse and of lexical change in which a huge part of the language involved in doing so is machine readable and in the public domain. Second, within the growing body of

---

[1]Retrieved 2013-09-11 from http://www.asexuality.org/home/about.html

research on language, gender, and sexuality, Kulick (1999, 2000, 2002, 2008) and Cameron and Kulick (2003, 2005) have argued for a need for more linguistic research on "desire", and among the numerous responses from scholars in the field (see Chapter 2 below), Eckert (2002) has argued that this should include linguistic research on "undesire." Through the study of asexuality, I will consider whether "undesire" is likely to be a useful analytical category for future research.

In addition, this dissertation makes novel contributions to corpus linguistic methodology. In generating a list of "keywords" that characterize a corpus of texts, a well-known problem is the need to exclude words extremely concentrated in a single part of the corpus (Gries, 2008). None of the commonly used approaches is appropriate for the forum data used in this dissertation (see Section 5.3 and Appendix B), and I propose using for this purpose a measure of dispersion used by Altmann et al. (2011).

In the rest of this chapter, I will briefly consider the first of these two issues (which will be discussed in more detail in Chapter 2). Then to contextualize the study of asexual discourse against broader social changes, I will discuss what, I argue, are three major factors in why asexual discourse has arisen when and how it has: widespread social use of the internet (Section 1.3.1), changes in marital and sexual ideologies during the 20th century (Section 1.3.2), and the social organization of homosexuality and gender variance (Section 1.3.3). I will then finish the chapter with an overview of the rest of this dissertation.

1.1   Lexical Change

A major contribution of variationist sociolinguistics has been to provide a means of linking synchronic and diachronic linguistics through quantitative analysis of structured heterogeneity Weinreich et al. (1968). While this paradigm has provided very fruitful in understanding the causes of sound changes and investigating ongoing sound changes, it has proven challenging to apply to other levels of linguistic analysis (for a recent overview, see Terkourafi, 2011).

On account of advances in computing technology, the examination of existing corpora as well as ones developed for a specific research project, new possibilities for understanding various aspects of lexical change have opened up, and this has included such diverse topics as maintenance vs. regularization of irregular English verbs Michel et al. (2011), the relationship between frequency of use and the rate of word replacement in Indo-European (Pagel et al., 2007), the lifespan of scientific terminology in public discourse (Bentley et al., 2012) and the roles of drift vs. selection in academic key words in different disciplines (Bentley, 2008). In addition to relying on relative frequencies (or related measures), some of the research made possible by large corpora investigate other measures, such as the dispersion of words over users and threads compared to what would be predicted by chance (Altmann et al., 2011). Through the use of a number of asexual corpora I have compiled (see Chapter 5 for details), I will investigate the development of asexual terminology over time, and this terminological evolution will be used to investigate changes in concepts over time. [2]

In examining developments in asexual discourse, a matter of relevance is that *discourse* is used in academic literature in, roughly, two main ways. One meaning refers to units of language above the sentence level (although even in this general sense, definitions vary; see Schiffrin, 1994, chapter2)), and the other is a larger, sociological set of language, beliefs, practices, etc. surrounding some concept. Sometimes these are distinguished by using discourse with a lower-case d for the former and Discourse with a capital-D for the latter (e.g. Gee, 1991). Because the "Discourse" of asexuality (i.e. the latter sense of *discourse* seems to have not existed prior to widespread use of the internet, examination of online asexual discourse (meaning the collection of the language, and sometimes images, used in asexual websites, blogs, forum threads, etc.) allows for a case study of the rise the of a new "Discourse," and how this relates to lexical development/change.

---

[2]This approach to examining conceptual evolution relies on the assumption that there is an important relationship between concepts and the lexicon. The relationship between the concepts in a given culture and the words in the language(s) used in that culture is a controversial matter (Pinker, 1994; Wierzbicka, 1997), but the assumption I am using is a less controversial one: there is a close connection between the terminology and the concepts in specific knowledge domains. Widespread acceptance of this is seen in the common practice for textbooks (including many linguistics textbooks) to put key terms in bold, and possibly include a glossary of these in the back. This relies on the assumption that learning the terminology for a specific domain is a particularly important part of learning the concepts.

In applying models of language change to asexual discourse, one particular model I intend to investigate is the "invisible hand" model Keller (1994), and its applicability to the development of asexual discourse. Keller applied to concept for an "invisible hand" (derived from economic theory) to language change to deal with problems stemming from trying to think of languages in terms of either a biological organism or as a human invention. The basic idea is to think of language neither as an organism (where human intention is irrelevant) nor as an invention (where its structure is determined by the intentions of its inventor), but as a "third kind" where structure arises through patterns of human behavior, with many people having parallel intentions, but where the changes are not generally the result of someone consciously trying to bring about those changes. (An example of such a "third kind" structure would be the development of walking paths on college campuses.) An important insight of this model is that to understand why language (or some other "third kind" structure) takes the shape that it does, its function must be considered (i.e. what do people use language for, and what motivations for using language in particular ways are shared by many users of a given language variety.) Because asexual discourse has developed in large part through people's intentions to create asexual communities and asexual discourse, I will consider whether an invisible hand model is an appropriate one for the development of online asexual discourse.

1.2 Language and Sexuality

Since the mid-1990s, there has been a growing body of work on language and sexuality at the intersection of sociolinguistics and linguistic anthropology. This research has developed out of work on language and sexual orientation on the one hand, and also out of research on language and gender on the other. A major debate within the field was sparked by work by Kulick (1999, 2000, 2008) and Cameron and Kulick (2003, 2005), who argued that much of the existing literature had too much of an emphasis of sexual identity, to the detriment of other aspects of sexuality. To remedy this, they recommended that more aspects of sexuality be studied, suggesting desire as one possible avenue for doing so. This invited

numerous responses from other scholars, one of whom Eckert (2002) argued that linguistic investigation of desire must also encompass the study of undesire. Drawing on the distinction between contraries and contradictories in antonyms Horn (1989) I will use *undesire* as the contrary of desire (i.e. desire to not X) and *nondesire* to mean its contradictory (no desire to X).

Within asexual discourse, the difference between *undesire* vs. *nondesire* is sometimes expressed in a distinction between asexuals who are "indifferent" (to sex) vs. the who are "repulsed". For example, these terms have been listed on the AVENwiki's "Lexicon" page since October 2010:[3]

> indifferent: 1) used by some asexual individuals to indicate that they feel neither revulsion toward nor powerful desire to engage in sex. 2) also can be taken to mean they are indifferent toward the idea of sex in general.
>
> repulsed: A term used by some asexual individuals to indicate that they find sex disgusting or revolting.

1.3 Historical factors necessary for the rise of asexual communities and discourse

Historical questions such as why asexual discourse arose when it did, where it did, and in what form(s) it did are causal questions where rigorous hypothesis-testing is probably impossible, but in my view, there are three especially important factors that can be asserted with reasonable confidence: widespread social use of the internet, a move in Western culture (which is also occurring in many non-Western countries) towards companionate marriage, and a certain social organization of homosexuality.

1.3.1 The internet

Early internet usage was limited to academic and government use. By the late 1980s and early 1990s, the technology and infrastructure–both for personal computers and for networks

---

[3]Retrieved 2016-05-29 from http://asexuality.org/wiki/index.php?title=Lexicon. The date they were added can be determined by the "history" page on the wiki.
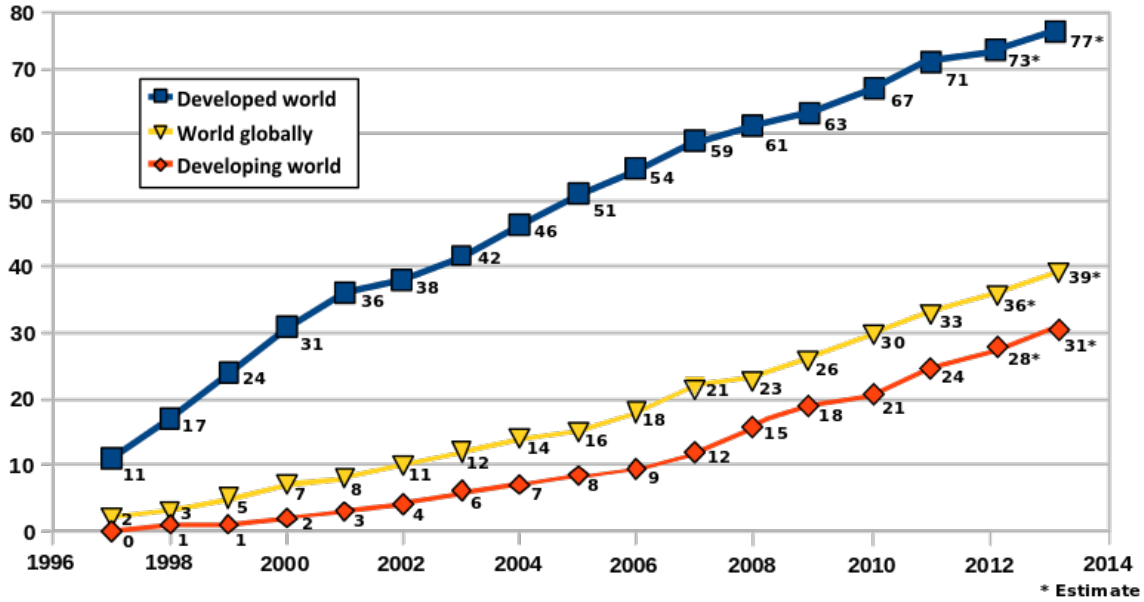
Figure 1.1: Internet users per 100 inhabitants, 1997-2011, with estimates for 2012-2013.(Wikimedia Commons, 2012)

that could link them–had developed far enough that there was beginning to be commercially available technology for the general population Leiner et al. (2009). Since then, the number of people in the world using the internet has grown enormously, as can be seen in Figure 1.1, which displays data on the percentage of people who use the internet in the developed world and in the developing world from 1997-2011. Internet use in developing countries has lagged far behind use in developed countries, but has been increasing.

A number of lines of evidence suggest that widespread use of the internet has been necessary for the development of asexual discourse. First, after considerable searching, I have been unable to find any pre-internet asexual organizations with any sort of continuity with online asexual communities. Because the strength of negative evidence depends on how thoroughly one has searched, in a survey for asexual-identified individuals 25 years of age or older (see Chapter 6 for details), I asked participants whether they participated in an offline asexual organization prior to finding an online asexual community, and likewise for asexuality-related organizations.[4] Of the 208 responses, only one person said yes to the for-

---

[4]The survey defined asexual communities/organizations as "ones that specifically used the term *asexual*, and it then defined

7

mer (a PTSD support group). For the question about offline asexuality-related communities, only eight said *yes*, and the majority of these were religious communities.

A second line of evidence comes from the existing historical accounts of asexual organizing. The most extensive accounts I know of are two by David Jay (2003, 2006), the founder of AVEN, one by Nat Titman (2012), the founder of the Asexuality community on LiveJournal, which for a while was the second largest, one on the AVENwiki by various authors,[5] and two by myself (Hinderliter, 2009a, 2013). All of these accounts begin with online communities. According to the AVENwiki, "Until the development of the internet, asexuals seldom had ways to connect with each other and share their experiences. In the early and mid 90s, people would occasionally post on sexuality-related newsgroups expressing their identification with asexuality or their lack of sexual attraction, but search functions were not effective enough, making it difficult to locate people with similar experiences."[6]

A third line of evidence is the sense that asexual organizing primarily occurs online, and that creating more offline asexual communities is a felt-need. For instance, a 2009 thread on the asexual forum Apositive called "Year in review" discusses highlights in asexual organizing during 2008, and one member responded by saying that their main hope for 2009 was to have more off-line asexual organizing. Two users responded with strong agreement, and none expressed disagreement. One expressing agreement was involved with asexual meet-ups in San Francisco and wrote that the previous meet-up had ten people, which was the most they had ever had. My impression at the time was that the San Francisco asexual meet-ups were among the most regular and best attended in the US. In 2012, the blogger Sciatrix wrote a post called, "An argument for offline organization (from a blogger on the internet)"[7] giving a number of reasons why such organizations are important. The title and the post suggest a sense that there are plenty of online locales, but not nearly enough off-line ones. Indeed, the post gives several reasons for why more offline organizing is needed, but does not argue

---

(on the same page) asexuality-related communities/organizations as ones "for issues related to asexuality (e.g. anti-sexual, celibate, abstinent, etc.)"

[5]Retrieved 2013-03-10 from http://www.asexuality.org/wiki/index.php?title=History

[6]Retrieved 2016-01-31 from http://www.asexuality.org/wiki/index.php?title=History

[7]Retrieved 2013-03-10 from http://asexualagenda.wordpress.com/2012/09/10/an-argument-for-offline-organization-from-a-blogger-on-the-internet/

that there was not very much at the time, presumably because readers were simply assumed to already agree with this.

These lines of evidence suggest that, if there was asexual organizing in Anglo-American countries prior to the formation of online asexual communities, they were small and had minimal influence on subsequent asexual organizing, and that the development of the internet was a necessary condition for the development of asexual discourse. However, it is not a sufficient reason. For asexual communities to form, people must have had some reason for wanting to participate in such communities. I suggest that the main factors for this are changing social attitudes about sex, marriage, and intimacy, specifically increased acceptance of pre-marital sex, an increased emphasis on companionship in justifying marriage, and an increased emphasis on sex as a part of that companionship.

1.3.2 Changes in marital and sexual ideologies

Bullough (1994) argues that over the course of the 20th century, the increasing social status of women and increasing demands for greater gender equity (including in marriage) made traditional understandings of marriage and sexuality less acceptable, and that one way of dealing with this "was to redefine marriage by emphasizing the sexual and affectional basis of intimacy and the importance of companionship in marriage" (1994: 138).[8]

Further, there was an enormous shift in sexual attitudes in the US during the 1960s and 1970s, which is popularly known as the "sexual revolution." Survey data collected during this time suggests that acceptance of premarital coitus in the US rose from about 20% of the population in 1963 to around 71% in 1978, and after this the rate more-or-less leveled off (Reiss, 2006). Various explanations have been given for this change, including the increasing role of women in the workforce (Reiss, 2006), the collapse of the syphilis epidemic following the discovery that it can be cured by penicillin (Francis, 2013), and availability of oral contraception (though Reiss and Francis both provide compelling data that, if a factor, it was not among the strongest).

---

[8]His main evidence for this change over time in popular opinion was changes over time in popular marriage manuals in the first half of the 20th century.

The importance of this in the creation of online asexual communities relates to Keller's (1994) application of an invisible hand theory to language change: Specifically, the importance of understanding the function(s) of online asexual discourse to understand how its structure has developed. To understand the function, it is important to know what felt-needs are being met. To date, there is no published academic work specifically about asexual communities, their dynamics, and people's reasons for participating in them, but using qualitative data from interviews and an online survey, Carrigan (2011, 471) identifies a number of common aspects in asexual-identity formation: "Adolescent experience gives rise to a sense of difference from a peer group, provoking self-questioning and the assumption of pathology (i.e. 'I thought there might be something wrong with me') before self-clarification is attained through the acquisition of a communal identity." While some individuals attained this self-clarification on their own, most in his study did not do so until finding an (online) asexual community. Presumably, this generally requires peers openly expressing sexual interest, or some sort of communication to enable the person to develop a concept of "normal sexuality" from which they diverge. Further, this perceived difference must be negatively valued, rather than (for instance) seen as virtuous.

From my own experience in online asexual communities, a major issue of concern in asexual discourse concerns how to develop emotionally satisfying relationships, given the perceived expectation that people's emotionally closest relationship should be with a romantic partner, and that sex should be an important part of that intimacy. While some asexuals probably felt "broken" on account of their asexuality prior the 1960s, it is likely social changes since then have exacerbated this. In addition, the available evidence suggests that the overwhelming majority of people participating in online asexual communities are women. The largest dataset of participants in online asexual communities comes from the Asexy Census run by Asexual Awareness Week 2011, getting a sample of n>3000 (Miller, 2012). The gender questions were poorly designed for data analysis, and individuals' gender identities were classified as female, female-ish, male, male-ish, and other. Female and female-ish accounted for 64% of the participants, while male and male-ish only accounted for

13%. Data from samples not recruited from asexual communities also suggests more female than male asexuals (Bogaert, 2004, 2013; Lucassen et al., 2011; Nurius, 1983) This suggests that changes in sexual expectations for women in particular during the 20th century play a major role in the rise of online asexual communities and discourse.

It is possible that these two factors alone–widespread social use of the internet, and changes in marital and sexual ideologies–would have been sufficient for the creation of online asexual communities, but I argue that a third social condition has played a major role in the specific forms that asexual discourse has taken: the social organization of homosexuality and gender variance.

1.3.3   Social organization of homosexuality and gender variance

Certain aspects of the social organization of homosexuality and gender variance in present day Anglo-America are historically quite unusual. In historical and cross-cultural work on homosexuality, there is some interest in classifying social organizations of homosexuality. Herdt (1991a, p.483) notes that, while typologies vary, they typically include at least four types: "age-structured, gender-transformed, class- or role-specialized, and modern gay/lesbian categories of homosexuality." In his first article, he focused on the first of these, and in the second, the second, and these appear to be the most common social organizations of homosexuality. A well-known examples of age-structured homosexuality is pederasty in Classical Greece(Herdt, 1991b). An example of (male) gender-transformed homosexuality is found in a number of countries surrounding the Mediterranean and in Latin America, where the primary distinction is not homosexual vs. heterosexual, but penetrator vs. penetrated, and men who penetrate other men in sex remain "men", while males who are penetrated in sex are grouped with women in a "not man" gender (c.f. Kulick, 2009)).

Donaldson and Dynes (1990) note that it is generally only in industrialized societies in Western Europe and North America (I would add "and cultures greatly influenced by these") where homosexuality is organized in a way where "both partners are adults and neither relinquishes his manhood or her womanhood, [and where] sexual reciprocation and

11

sexual role reversal are generally honored if not universally practiced, and in theory the partners are equal."

An important part of the social organization of homosexuality and gender variance in the US today is that they are treated as separate categories. Further, men predominantly/exclusively attracted to men and women predominantly/exclusively attracted to women are grouped together in the category "homosexuality." In recent decades, LGBT (lesbian, gay, bisexual, transgender) or some variation of the initialism, often with additional letters, has become a major part of activism in this area. This involves an often tenuous alliance between groups with some shared interests and concerns. Further, homosexuality and gender variance are increasingly approved of in US society at large (e.g. attitudes about gay marriage have changed considerably over the past 15 years or so (Pew Research Center, 2015), and this is by far the dominant position in some parts of society (e.g. in academia and the mental health professions). As such, they are politically well-organized enough to make useful allies for asexuality, and are liked well-enough for alliances to not pose a serious risk to asexual politics.

LGBT groups have been the primary groups that asexual activism has tried to build connections with, and this is proving to have a fair amount of success. Within asexual communities an issue that is sometimes debated is sometimes called the "LGBT inclusion" issue–the question is whether asexuality is/should be included under the LGBT umbrella, although those most involved in asexual activism are overwhelmingly on the pro-inclusion side, or at least not opposed to it (p.c. Michael Doré).

1.4 Overview of the rest of this dissertation

The structure of this dissertation is as follows. In Chapter 2, I review previous work on language and sexuality (Chapter 2.1), and on lexical change (Chapter 2.2). Chapters 3 and 4 contextualize and motivate the research questions for the quantitative parts of this study: Chapter 3 examines historical and relatively recent academic discourse concerning

asexuality, and then Chapter 4 explores the history of asexuality, beginning with what can be known about the decades prior to the formation on online asexual communities, about early online asexual discourse and communities. This chapter explains many of the terms and concepts explored in later chapters.

Chapter 4 proposes that there have been (at least) two major conceptual shifts in the conceptualization of asexuality in the time period under consideration. The first, which I am calling the *AVEN shift* (Chapter 4.5.1) is that the conceptualization of asexuality in asexual discourse converged, to a large extent, in the direction of certain parts of how AVEN presents asexuality. (AVEN's presentation of asexuality also changed in cases where concepts were not widely adopted.) Second, there has been a *rise of intermediate categories* (Chapter 4.5.2): There has been an increase in usage of convergence on terms that refer to people who are, in some sense, in between being asexual and not being asexual, and these terms increased in importance in asexual discourse. Chapter 4 provides primarily qualitative evidence for these, motivating quantitative investigation in Chapters 5 and 6.

Chapter 5 uses corpus data to explore characteristics of online asexual discourse. First, I ask whether online asexual discourse can be reasonably well-distinguished from similar kinds of discourse. Second, if so, what terms are most distinctive. Third, does data about usage over time support the *AVEN shift* and *rise of intermediate categories* hypotheses? (e.g. are the words identified as being relevant especially prominent features of online asexual discourse?)

To answer these questions, I constructed four corpora (Chapter 5.1): I scraped the largest asexual website (Asexual Visibility and Education Network [AVEN]) and a similar sized message board on a different topic as a control (*Non-asexual corpus*). I subdivided the AVEN data into two sub-corpora, based on the sub-forum topics. Most subforums are in the *AVEN-core Corpus.* Some (e.g. "Just For Fun" or "Off-A") are mostly about topics other than asexuality, and were grouped as the *AVEN-other Corpus.* In addition, I scraped several asexual blogs and asexual communities other than AVEN (e.g. a LiveJournal community): these comprise the *Asexual-other corpus.*

In Chapter 5.2, I describe the results of a multinomial Naïve Bayes trained on the AVEN-core and reference corpora. I found moderate distinguishability between the AVEN-main and non-asexual corpora at the level of individual posts, when only considering individual words. To rule out the possibility that the classifier was distinguishing AVEN vs. non-AVEN discourse, I first created an algorithm to remove from consideration words that probably refer to users (Appendix A). Second, I used the same classifier on the AVEN-other corpora and the asexual-other corpora. Results for the Asexual-other discourse are similar to result for the AVEN-main, while results for the AVEN-other corpus do not. This suggests that the classifier is identifying asexual discourse vs. other discourse.

In Chapter 5.3, I use a modified version of the metrics used by the classifier to provide a list of key-words that are characteristic of asexual discourse. Because the text in the corpus data is time-stamped, I use the AVEN-core and Asexual-other corpora to examine the usage over time for three sets of words that appeared among the keywords: intermediate category terms, romantic orientation terms, and terms relating to the *repulsed* vs. *indifferent* distinction. I find that many terms relevant to the hypotheses of Chapter 4 are indeed characteristic of asexual discourse. And I find strong evidence to support the "rise of intermediate categories" hypothesis.

However, it is harder to test the "AVEN shift" hypothesis with the corpus data, because it largely involves changes in the time period when the data is the most sparse, and because it involves people's views changing before vs. after encountering online asexual discourse. Because of this, Chapter 6 reports the results of a survey (using retrospective reporting) to further investigate the "AVEN shift" hypothesis. Results generally support the hypothesis that there has been a shift towards AVEN's presentation of asexuality, although the change is by no means monolithic.

Chapter 7 concludes this dissertation, suggests avenues for further research on online asexual discourse, and suggests how the methodologies used here can be applied to study the evolution of discourse in other online communities.

Chapter 2

Survey of current research

Although there has been a growing body of research on asexuality, most of this research has been published within the past decade and has primarily come from psychology (e.g. Bogaert, 2004, 2006; Brotto et al., 2010; Brotto and Yule, 2011; Prause and Graham, 2007), sociology (e.g. Carrigan, 2011; Scherrer, 2008; Poston and Baumle, 2010), and gender studies and closely related fields (e.g. Kim, 2010; Cerankowski and Milks, 2010; Przybylo, 2011). To date, none of the published research on asexuality is from a linguistic or anthropological perspective,[1] but there is a growing body of literature on language and sexuality, largely at the intersection of sociolinguistics and linguistic anthropology.

This dissertation primarily addresses two general issues within linguistics, one concerning on language and sexuality, and one concerning language change and the development of a Discourse (capital-D). The first of these will be discussed in Section 2.1 and the second in Section 2.2.

2.1 Language, sexuality, and asexuality

The study of language and sexuality largely finds its roots in earlier research on language and gender and in research on language and sexual orientation. In Section 2.1.1, the shift from "language and gender" to "language, gender, and sexuality" will be discussed. In the literature on language and sexuality, there have been claims of a move away from a lexical focus, and Section 2.1.2 will survey lexically focused research concerning sexuality, and I argue that my lexical focus is consistent with these trends in lexically-focused work in

---

[1] I have published some theoretical/methodological pieces Hinderliter (2009b, 2013), but these were primarily addressed to non-linguists.

this area. In addition, my research on asexuality addresses an ongoing debate in the field concerning identity and desire as analytical tools and emphases in research, and relevant details of this debate will be reviewed in Section 2.1.3.

2.1.1 Language, gender and sexuality

Much of the research on language and gender in linguistics is generally traced back to threads from the 1970s, although sometimes earlier work (e.g. Jespersen, 1922; Sapir, 1992) is referred to. The first of these threads is the use of sex/gender as a variable in variationist sociolinguistic research. The second of these was the controversy and research that followed from Lakoff's (1973) article "Language and Woman's Place." The ensuing research (as well as some earlier work) is often described with three D's (e.g. Baker, 2008; Cameron, 2005): the deficit approach, the difference approach, and the dominance approach, with the latter of two becoming the two primary models in the field. Both focused on differences between men's and women's speech (as did the variationist approaches), but they tended to differ in explanatory mechanisms. The dominance approach saw gender differences in language use as stemming from male dominance, whereas the difference approach focused on differential socialization and the strong tendency for children to be socialized in same-gender groups.

Later work criticized all of these strands in that they focused on the difference between men and women, but not on the variation among women (or among men). As Cameron and Kulick (2005, p.582) states when describing a broad shift in the field of language and gender, "Since the first half of the 1990s, a paradigm organized around the concept of a binary gender difference has been superseded...by one that is concerned with *diversity* of gender identities and gender practices" (italics original). Part of this shift involves a move away from a mainstream focus to a more liminal focus, which they describe as "more interest in non-mainstream and 'queer' gender identities, and in [the] relation of gender and sexual identities to heteronormativity" (2005: 484). This focus on the liminal is driven by various factors, including academic activism concerning the studied groups as well as a belief that a focus on the marginal will help to shed light on issues of power and ideology

of the mainstream. A part of this shift towards the liminal has involved research on sexual minorities, thus moving the subfield of language and gender in the direction of language, gender, and sexuality. In the next two subsections, I will discuss two strands in the literature on language and sexuality, especially sexual orientation. The first is more methodological and concerns research focusing on the lexicon, and the second is more conceptual and involves a debate about use of identity and desire as analytical perspectives.

2.1.2   Lexical research in the language and sexuality literature

A point made by many reviews of the literature on language and sexual orientation (e.g. Cameron and Kulick, 2003; Kulick, 2000; Livia and Hall, 1997a; Queen, 2007) is that most of the early research had a lexical focus, specifically looking at lexical items that were ideologically linked to male homosexual subcultures. This emphasis–and a desire to move away from it–is clearly seen in the title of one of the first edited volumes and gay and lesbian language: *Beyond the Lavender Lexicon* (Leap, 1995). In a recent survey of the field, Leap and Motschenbacher (2012), suggest that, "Studies on vocabulary...have been a source of insights into the interface of language and sexuality," but for references they mention "early contributions on 'homosexual argot and slang'" in Cameron and Kulick's (2006) language and sexuality reader and then go on to suggest that research has increasingly moved towards a focus on stylization.

Nevertheless, research with a lexical focus has continued to be an important part of research in this area. Even in *Beyond the Lavender Lexicon*, several chapters have a lexical focus. Porter (1995) examines Kenyan discourses about homosexuality, with considerable attention to various lexical items that help to illustrate the local ideologies and ways of categorizing and criticizing people. Bolton (1995) engages in quantitative corpus analysis of gay male erotica, with the primary data analysis focusing on the frequency of references to specific body parts and modifiers used to describe these, especially for the penis and semen. Manalansan (1995) examines language use by Filipino gay men living in the US, with particular attention to terms for categorizing people and terms he found especially

salient and helpful for understanding cultural meanings and the tension between a Filipino form of homosexuality based on a penetrator/penetrated distinction currently in competition with the organization of homosexuality now dominant in the US focusing on a heterosexual/homosexual distinction.

Probably the next two major edited volumes in the literature on language and sexuality are *Queerly Phrased: Language, Gender and Sexuality* (Livia and Hall, 1997b) and *Language and Sexuality: Contesting Meaning in Theory and Practice* (Campbell-Kibler, 2002). About a third of the chapters in *Queerly Phrased* are in a section called "liminal lexicality," which includes descriptive work on terms relating to homosexuality (and sometimes bisexuality) in Japanese (Valentine, 1997), American Sign Language (Kleinfeld and Warner, 1997), pre-modern French (Conner, 1997), and 16th century English (Watt, 1997), among other topics. In addition to basic descriptive work, (Murphy, 1997) reports on survey data testing social psychological theories concerning social categorization applied to bisexuality (Murphy, 1997).

Likewise, several chapters in Campbell-Kibler (2002) have a lexical focus. Livia (2002) examines personal ads in a French magazine for lesbians, and part of her analysis focuses on frequently used terms for who is told not to apply, and also terms for what traits are considered desirable; along with editorial comments; this functions as part of an analysis of the ideology the magazine is trying to push. In earlier work, Wong and Zhang (2000) focused on the Chinese word *tongzhi*, which had been used in communist politics to mean comrade but fell out of favor in the aftermath of the Cultural Revolution; it was later adopted by many gay men and lesbians in Taiwan and Hong Kong as a term to refer to themselves. Wong and Zhang use this term as a means for examining resources for constructing an imagined Chinese gay community, with its own distinct style. Wong (2002)further examines *tongzhi* by considering the role of Hong Kong newspapers in its semantic derogation within the general population.

Two monographs with a considerable amount of lexical focus are Baker's (2002) work on Polari, a counter-language used by gay men and people in the theater in the UK in the earlier part of the 20th century, and Baker's (2005) book, *Public Discourses of Gay Men.* That

work aims to demonstrate the utility of corpus linguistic methods for studying language and sexuality, especially to better understanding (often competing) ideologies in discourse. He does this by applying various corpus approaches to examine issues such as a House of Lords debate about whether to make the age of consent for same-sex sexuality the same as for heterosexuality, comparing gay and lesbian erotic fiction, safer sex publications, the show *Will and Grace*, and discussion of homosexuality in British tabloids. For each topic, he had to compile a corpus, which he did not grammatically tag, and so a lexical focus was often necessary as a matter of quantification.

In this lexically-focused research since the mid-1990s, two main aspects stand out. First, most of it either focuses on understudied populations (i.e. ones other than present day English-speaking gay men and lesbians) or it involves quantitative corpus or survey research. Second, the focus on the lexicon is not simply a list of words in some subculture, but the terms highlighted are used to illuminate beliefs, ideology, and/or some kind of social organization.

My own research on asexuality will have a largely lexical focus, and this is consistent with recent lexical research in the field of language and sexuality: asexuals are an understudied population, and a lexical focus is used for more than just a glossary of words. Further, my use of corpus linguistic methods makes an emphasis on the lexicon of practical value, and my use of corpus linguistic methods and a survey will enable me to move beyond basic description.

2.1.3 Language, identity, and desire

Part of the move away from a lexical focus in the research on language and sexuality concerns a reaction against much of the earlier literature and attempts to find some kind of authentic gay language, with lexical items being among the most salient aspects of the supposed gay language. The reaction largely stems from a desire to avoid any sort of "essentialist" understanding of identity, and a major debate in the literature involves how to do this, with controversy centering on the use of desire and identity in language and sexuality research.

In a number of publications, Kulick (1999, 2000, 2002, 2008) and Cameron and Kulick

(2003, 2005) have argued that much of the then existing research on language and sexuality had an excessive emphasis on sexual identity, and they argued that additional emphases were necessary. As one possibility, they offer "desire" as a theoretical perspective for research to focus on, suggesting a psychoanalytic framework for doing so. While they avoid endorsing any of the frameworks of desire they examine, they do suggest a number of themes: sexual desire is not entirely conscious or rational, it is unavoidably both social and linguistic, it involves both a desiring subject and a desired object, that not all desire is sexual, and that any analysis of desire should consider the power relations involved in allowing or preventing its expression. To illustrate what desire-oriented research might look like, they draw attention to a number of existing studies that could be framed in terms of desire.

Cameron and Kulick's sometimes provocative suggestions have received considerable response from other scholars (e.g. Barrett, 2002; Bucholtz and Hall, 2004; Eckert, 2002; Queen, 2007; Wong et al., 2002), with the result that issues of identity and desire have become a major site of theoretical controversy in the field Morrish and Sauntson (2007); Queen (2007). In Eckert's (2002), she recognizes the problems this proposal is trying to deal with, but expresses a degree of hesitancy and caution towards the proposal to shift focus to desire. Because people tend to think of desire as something more internal (and therefore not directly observable) than activity–and it is the activity aspect of sexuality that makes it social–Eckert (2002, p.100) expresses fear that a focus on desire "brings a mystification into the study of sexuality, and it is in contemplating desire that we are inclined to fall into an asocial and naturalized view of sexuality." She therefore suggests means of studying the social mediation of desire, discusses some pitfalls to avoid, and considers some ways that sexuality is about much more than desire: "Sexuality is not just about desire; it is also about undesire. For many, perhaps particularly victims of assault, abuse, and unwanted attention, sexuality can be as much about revulsion, fear, or lack of desire as it is about desire" (p. 104).

Eckert did not expand much on undesire, and in this dissertation I use the investigation of online asexual discourse to explore the (potential) utility of undesire for understanding language and sexuality. As a first modification to her proposal, I suggest it is important–at

least for studying online asexual discourse–to employ the distinction between contraries and contradictories in antonyms and other expressions of negation, and so I will investigate the utility of both undesire (contrary, i.e. "want to not X") and nondesire (contradictory, i.e. "not want to X") as theoretical lenses for understanding online asexual discourse.

2.2 The development of a discourse and lexical change

Because asexual discourse, in something like its present form, has only been possible on any socially noticeable scale through widespread usage of the internet, much of the discourse data involved in constructing this Discourse remains available, is machine readable and in the public domain. Because a Discourse is a rather complex thing that is not directly measurable, more indirect methods are needed to examine developments over time. A challenge that qualitative approaches must deal with is that, at any given time in virtually any community, there is a diversity of opinion, and so simply analyzing two texts from two times cannot constitute evidence of change. One qualitative approach that avoids this problem is used by Blommaert (2005): To examine ideological changes over time in the Flemish socialist party, he examines and compares the preparatory texts for their 'ideological congress' in 1998 and the one for the previous such meeting, which was in 1974. A similar approach with asexual communities would be to look at changes over time in some online community's introductory material. Likewise, this kind of method works well with changes to the AVENwiki because of the way of that wikis keep a record to allow users to see a snapshot of the page after each edit.

A more quantitative approach is suggested by Baker (2005). To demonstrate the utility of using quantitative approaches from corpus linguistics to better understanding (oftentimes competing) ideologies, he employs various applications of using relative frequencies and collocations for examining various discourses concerning homosexuality in the UK, including changes over time. I adapt this approach to studying asexuality in Chapter 5

Chapter 3

Academic discourses about asexuality

In order to explore the origins and historical development of asexual discourse, it is prudent to investigate via multiple approaches. In Chapters 3 and 4 I rely primarily on qualitative approaches to understand the historical background in which asexual discourse arose and the development of that discourse. This chapter investigates academic discourses about "asexuality" (broadly and vaguely defined) and the next chapter will outline the history of online asexual communities, and the history of some important terminology.

While the history of online asexual communities is more directly relevant to conceptual and lexical developments in online asexual discourse, examination of academic discourses is important for three reasons. First, there have been significant interactions between academic discourse and online asexual discourse as will be seen several times below, and so familiarity with certain parts of the academic discourse is important for understanding the emergence of online asexual discourse. Second, some scholars are prone towards making sweeping claims about the newness of asexual discourse, or even of asexuality itself (e.g. Przybylo, 2013, says "Asexuality has not existed at any other time in Western history, not as 'asexuality' per se."). By examination of some major academic discourses since the late 19th century, we can see various attempts at addressing similar issues to those which are addressed in online asexual discourse.

The third reason involves a tendency in many modern discourses on the history of sexuality, where much has been made of an oft cited passage of Foucault's (1995):

> As defined by the ancient civil or canonical codes, sodomy was a category of forbidden acts; their perpetrator was nothing more than the juridical subject of them. The nineteenth-century homosexual became a personage, a past, a case history,

and a childhood... It was consubstantial with him, less as a habitual sin than as a singular nature. We must not forget that the psychological, psychiatric, medical category of homosexuality was constituted from the moment it was characterized— Westphal's famous article of the 1870s on 'contrary sexual sensations' can stand as its date of birth...The sodomite had been a temporary aberration; the homosexual was now a species. (p. 43)

In this telling of history, (the category of) "the homosexual" came into being with its first use in medical/psychiatric/psychological discourse. A great deal of work on the (modern) history of sexuality, drawing on Foucault, has emphasized the power and influence of psychiatric, psychological, and medical discourses. Some social constructionist work on asexuality attempts to force the history of asexuality into a narrative in which it is first constructed by medical/scientific discourses, and then this medicalization resisted/questioned/challenged. For instance, in an article titled "Asexuality: From pathology to identity and beyond," Gressgård (2013) attempts this. She draws on references to "the frigid wife" and "the impotent husband" in Foucault (1995, p.110) to establish a psychiatric construction of asexuality, and then discusses more recent academic work (e.g. Bogaert, 2006; Prause and Graham, 2007) as evidence of a more recent shift away from pathology towards identity (by which she means self-identification) in academic discourses about asexuality. Similarly, in an article about scientific discourses about asexuality, quoted above, Przybylo (2013) writes that "asexuality, like most sexualities, is in significant and intricate ways carved into existence by science," although she then goes on to qualify this by noting that asexual-identified individuals and various media pieces have also played a crucial role.

Given this academic backdrop, it is reasonable to examine academic discourses about asexuality in investigating the development of online asexual discourse. The rest of this chapter is organized as follows: In Section 3.1, I will examine conceptual options available in academic discourses for understanding asexuality prior to 2000. In Section 3.2, I will briefly examine more recent academic work that has coincided with the growth on online asexual communities.

3.1 Asexuality in sexological discourses before 2000

Within sexological discourses before 2000, I have found four main ways in which asexuality–which I use in this section in an intentionally vague way, hopefully without risk of anachronism–has been treated: as pathology, as a throw-away category, as a sexual orientation category, and as preferential celibacy.

3.1.1 Asexuality as Pathology

Richard von Krafft-Ebing (1998), in his sexological classic *Psychopathia Sexuallis*, treats asexuality as pathology. His methodology consists primarily in giving case histories of patients and making taxonomies of sexual disorders. Within his classification, he lists four kinds of *cerebral* neuroses": *Paradoxia* (sexual excitement in childhood or sexual instinct reawakening in old age), *anaesthesia* (lack of sexual instinct), *hyperaesthesia* (excessive sexual desire), and *paraethesia* (perversion of sexual instinct). His work is best known for his discussions of the "sexual perversions" (paraethesia), which is what the bulk of the book is about, but of relevance to asexuality is his discussion of anaesthesia, the subject of his case studies 3-10. He divided anaesthesia sexualis into two types: "as a congenital anomaly" and "acquired anaesthesia." The former are cases where "in spite of generative organs normally developed and the performance of their functions (secretion of semen, menstruation), the corresponding emotions of sexual life are absolutely wanting. These functionally sexless individuals are rare cases" (Krafft-Ebing, 1998, p.40). One of the case studies (case study 8) is of a woman who had never experienced libido and never felt sexual excitement in coitus with her husband. All other case studies were of men. These include four "genuine cases of anesthesia", which are about men who had never masturbated and never (or virtually never) felt sexual excitement, except perhaps in dreams. Some had tried coitus and were either unable to maintain erection or found it unexciting.

Krafft-Ebing (1998, p.42) also includes two cases of men in whom "the mental side of the sex life is a blank leaf in the life of the individual, but where elementary sexual sensa-

tions manifest themselves at least in masturbation." Neither of these men was interested in coitus, but both masturbated to relieve sexual desire. The second is described as having an abhorrence of sex and a hatred of women–especially those who had been pregnant, as he saw reproduction as the cause of human misery. Krafft-Ebbing also describes a transitional case between this condition and "genuine anaesthesia sexualis." Krafft-Ebbing gives no case studies for acquired anaesthesia sexualis, but lists a number of alleged causes for loss of sexual desire. Reading his work against current asexual typologies, two important issues of classification are already present: how to group "asexuals" who masturbate versus those who do not, and whether to group with "asexual" extreme negative views of sex (in current asexual discourse, this is labeled antisexuality).

Within sexology, pathology remains a conceptual lens for understanding lack of sexual desire.[1] The primary classification of mental disorders in the United States is the Diagnostic and Statistical Manual of Mental Disorder (DSM), published by the American Psychiatric Association. In DSM-I (American Psychiatric Association, 1952), no sexual dysfunctions were listed, but there was a section about psychophysiological genitourinary disorders. DSM-II (American Psychiatric Association, 1968), added impotence and dyspareunia among the examples of this. Masters and Johnson's (1970) book *Human Sexual Inadequacy* listed several sexual dysfunctions, which mostly concerned arousal, orgasm, ejaculation, and pain. Their work was instrumental in the subsequent development of sex therapy in the 1970s (Irvine, 2005). In 1977, two sex-therapists, Helen Singer Kaplan (1977) and Harold Lief (1977) proposed a disorder based on lack of interest in sexual activity (or in the right kind of sexual activity). Consistent with other sexual dysfunctions they proposed a subtyping system in which the dysfunction can be lifelong/primary or acquired/secondary, and it can be situational or generalized.[2]

Using Lief's term, Inhibited Sexual Desire (ISD), DSM-III (American Psychiatric Asso-

---

[1]The following is largely adapted from a fuller account by Hinderliter (2013).

[2]Lifelong (or primary) is for when the person has always been uninterested in (the right kind of) sex since puberty, whereas acquired (or secondary) indicates that it was subsequent to a period of normal sexual functioning. The condition is generalized if the person is not interested in sexual activities with anyone (or alone), while it is situational if the lack of interest is restricted to their current partner (or appropriate partners more generally).

ciation, 1980) included this in its newly added section of sexual dysfunctions. In DSM-III-R (American Psychiatric Association, 1987)[3] ISD was replaced with two diagnoses: Hypoactive Sexual Desire Disorder (HSDD) and Sexual Aversion Disorder (SAD). Of these two, HSDD has received more attention, and so I will focus more on it, as have other authors discussing asexuality and sexual dysfunctions. Since the publication of DSM-IV (American Psychiatric Association, 1994) these diagnoses have had a criterion (called clinical significance criteria) indicating that to be diagnosed, it must cause "marked distress or interpersonal difficulties" (American Psychiatric Association, 2000, e.g. p.541). Similar requirements were added to the diagnostic criteria of some diagnoses in DSM-III-R and a great many in DSM-IV, such that over half of the diagnoses in DSM-IV have clinical significance criteria (Spitzer and Wakefield, 1999). While life-long HSDD is conceptually similar to asexuality, this conceptualization has resulted in minimal research, despite being in the DSM for over 30 years: I have not been able to find a single study specifically on life-long HSDD.[4]

3.1.2   Asexuality as a Throw-Away Category

Another way that asexuality sometimes functions in sexological work is as a throw-away category. Probably the classic case of this is seen in the two Kinsey reports (Kinsey et al., 1948, 1953). In the male volume (Kinsey et al., 1948), Kinsey proposes a 7-point bipolar scale for sexual orientation ranging from 0 (exclusively heterosexual) to 6 (exclusively homosexual), justifying it in a much cited passage:

> Males do not represent two discrete populations, heterosexual and homosexual. The world is not divided into sheep and goats. Not all things are black nor all things white. It is a fundamental of taxonomy that nature rarely deals with discrete categories. Only the human mind invents categories and tries to force facts into

---

[3]It had been decided that new editions of the DSM would be published to roughly coincide with new editions of the International Classification of Disease (ICD), but after DSM-III was published, it was decided that a new version of the DSM was needed before the next edition of the ICD, and so DSM-III-R (R means revision) was published in the interim. Many felt that this was too much change too fast, and so DSM-IV took a much more conservative approach to change, where all new changes other than the addition of clinical significance criteria required considerable empirical support.

[4]In addition to using pubmed, google scholar, and reading a number of chapters and articles summarizing research on HSDD, I sent out an email to SEXNET in October 2011, a listserv for sex researchers from a number of disciplines, asking if anyone knew of any studies on the matter. The only response I received was from the chair of the DSM-5 Sexual and Gender Identity Disorders workgroup, saying he was forwarding my inquiry to the two members of that workgroup responsible for the literature reviews for HSDD and Female Sexual Arousal Disorder. The only research specifically on life-long HSDD either knew of was a study that one of them was then conducting.

separated pigeon-holes. The living world is a continuum in each and every one of its aspects. The sooner we learn this concerning human sexual behavior the sooner we shall reach a sound understanding of the realities of sex (1948: 639).

Their operationalization is less than clear, but seems to have involved a combination of psychological factors and behavioral history. This approach fails to have any good way of dealing with asexuality and a residual category X was employed. The throw-away character of this category is perhaps best seen in an editorial oversight: Table 141 says, "For an explanation of the meanings of the ratings X, 0, 1, etc., see the accompanying text" (1948, p. 640), and similar or identical wording is found in five other tables. But there is no explanation for X in the accompanying text. Rather, the only place in the book where X is defined is in the caption of Figure 169: "Percent shown as 'X' have no socio-sexual contacts or reactions" (1948. p. 656). Because of this definition, if someone ever has sex with someone else, that person ceases to be X and are on the scale. This logic is applied in the only place where X is not entirely a throw-away category in the male volume, which are the tables showing distribution on the Kinsey Scale based on age, showing that the percent identified X "rapidly disappear between the ages of 5 and 20" (1948, p. 658). The low percent of adult men labeled X is likely why this category received so little attention. Kinsey's female volume (Kinsey et al., 1953, p.472) gives more attention to X, and the explanation for X seems to justify this:

> Individuals are rated as X's if they do not respond erotically to either heterosexual or homosexual stimuli, and do not have overt contact with individuals of either sex in which there is evidence of any response. After early adolescence there are very few males in this classification..., but a goodly number of females belong to this category in every age group.... It is not impossible that of some of these individuals might show that they do sometimes respond to socio-sexual stimuli, but they are unresponsive and inexperienced as far as it is possible to determine by any ordinary means.

The high rates of X for some categories (i.e. 17% for 20-year-old single females) suggests that Kinsey's operationalization was better suited for men than for women, and the last sentence suggests that his goal with X was to get as few people in it as possible.

Another way that asexuality functions as a throw-away category in sexological literature

is when it unexpectedly appears in the data-set, and researchers choose to throw out the data for asexuals, for instance when people enter asexual as a write-in option on a sexual orientation question (e.g. Lever et al., 1992; Vrangalova and Savin-Williams, 2012)

3.1.3 Asexuality as a sexual orientation category

A third conceptualization of asexuality in sexological discourse has been as a sexual orientation category. Although similar models were proposed by Magnus Hirschfeld (1896) [5] and by Shively and De Cecco (1977), the most cited version of this is one by Storms (1979, 1980), which seems to have been proposed independently of these others. As with Kinsey's bipolar scale, early research on masculinity and femininity had placed masculinity and femininity on a bipolar scale, with masculinity at one end and femininity at the other. In the 1970s, some gender researchers (Spence et al., 1975) criticized this model because it (wrongly, they argued) predicted that if someone was more feminine this must mean that that person was less masculine and vise-versa; this model consequently failed to distinguish between individuals who were rated high on both masculine and feminine traits and individuals who rated low on both. As an alternative to a bipolar scale, a model had been proposed with masculinity and femininity on two separate axes. On this basis, Storms suspected that Kinsey's model would have similar problems, failing to distinguish between bisexuals and asexuals. He thus proposed a two-dimensional model that placed heteroeroticism and homoeroticism on separate axes, thereby allowing for asexuality as a logical possibility, as seen in Figure 3.1.

Including asexuality in a model of sexual orientation does not necessarily mean that asexuality is thus seen as a sexual orientation. Some authors (e.g. Brotto et al., 2010; Brotto and Yule, 2011) interpret Storms's 1980 paper as treating asexuality as a lack of sexual orientation. In fact, Storms did not call asexuality a sexual orientation or a lack of sexual orientation, but simply said that his model generates four "sexual orientation categories."

---

[5]There does not appear to be an English translation of the book as a whole, a scholarly description is given by Haeberle (1998). In early 2014, An AVENite become interested in what Hirschfeld said about asexuality, and translated some relevant portions of *Sappho und Socrates* and other texts into English, and posted them to the AVEN website. Retrieved 2014-03-06 from http://www.asexuality.org/en/topic/98639-indirect-mentions-of-asexuality-in-magnus-hirschfelds-books/
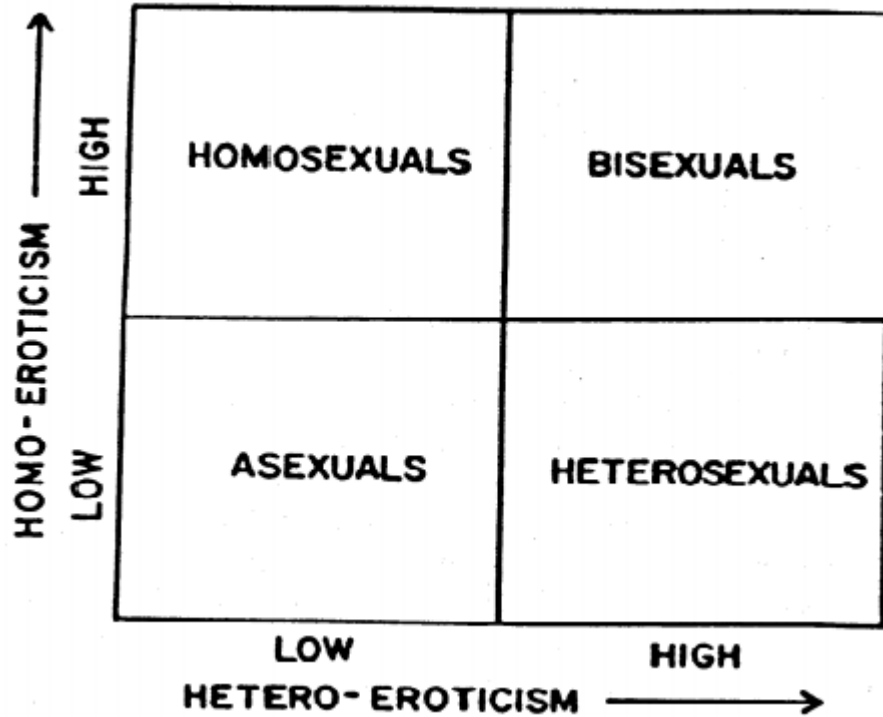
Figure 3.1: Storms's (1980) model of sexual orientation. Caption: "Four orientation categories generated by a two-dimensional model."

His avoidance of the term "sexual orientations" may be because he viewed asexuality as a lack of sexual orientation, but I suspect the question had occurred to him of whether asexuality is a sexual orientation or a lack of sexual orientation, and by calling it a "sexual orientation category," he dodged the question.

In addition to the word asexual, Storms (1979) also used the term "anerotic" to refer to the same group, though he did not use the word in the more often cited 1980 paper. In the pre-2000 sexological literature, one of the places where asexual as a sexual orientation category was most commonly used has been in literature using sexual orientation in typologies of male-to-female transsexuals–for instance, the word asexual is used in a list of sexual orientations in DSM-III's (American Psychiatric Association, 1980) diagnosis Transsexualism. Another term sometimes used in that context alongside the sexual orientation terms homosexual, heterosexual, and bisexual is "analloerotic,"[6] which was coined by Blanchard (1989a,b), who

---

[6]In psychoanalytic literature, alloeroticism involves sexual interest in or behavior with others. The term contrasts with autoeroticism.

restricted the term "asexual" for the subset that does not display any erotic interests (i.e. neither partnered nor solitary).

Despite the intuitive appeal of Storms' model over a bipolar one like Kinsey's, I have only been able to find five times such a model has been used (Berkey et al., 1990; Blanchard, 1985; Conner, 1996; Nurius, 1983; Nuttbrock et al., 2011), and bipolar models continue to be much more commonly used.

3.1.4 Asexuality as preferential celibacy

In asexual discourses (both academic and lay) the earliest known scholarly article specifically about asexuality is Johnson's (1977) chapter "Asexual and autoerotic women: Two invisible groups" appearing in a book titled *The Sexually Oppressed* (Gochros and Gochros, 1977). This article anticipated a number of arguments later made in online asexual discourse. Johnson opens by considering the language available in English for talking about such women, noting that when asexual is (infrequently) used in the scientific literature on human sexuality its meaning is generally vague. She then cites an example from popular usage where a character is described as asexual where this connotes "manipulated goodness." Johnson notes (using a dictionary for evidence) that "this connotation ... is also evoked by such synonyms as celibate, chaste, and virgin–all of which are identified, historically, with religious principles" (1977: 96-97). She then notes a problem of vocabulary that will be seen below when early asexual websites are examined. After commenting that these terms suggest that sexual desire is restrained rather than non-existent, she writes:

> There appear to be very few really appropriate words in the English language to describe the individual who, regardless of physical or emotional condition, actual sexual history, and marital status or ideological orientation, seems to prefer not to engage in sexual activity. Oppressed by consensus that they are nonexistent, these are the 'unnoticed' who in this article are called 'asexual'–by default. (1977: 97, italics original)

She later defines "autoerotic women" in a way somewhat inconsistent with this definition: "While the asexual woman, who has no sexual desires at all, is almost completely

unrecognized, the autoerotic woman, who recognizes such desires but prefers to satisfy them alone, is similarly dismissed" (1977, p.99). Using letters to the editor in women's magazines as her data, she discusses how the women in these groups have had "their personal experiences...redefined for them in terms of socially constructed meanings" (1977, p. 104), focusing on four: as ascetic, as neurotic (i.e. what am calling "pathology"), as unliberated, and as politically conscious (i.e.having their sexuality co-opted as an expression of feminist separatism).

Johnson's analysis makes a distinction between those who masturbate and those who do not. Yet this distinction is inconsistent with her first definition which would include both groups, and the letters to the editor she examines provide no data about which of these categories their authors belong to. The conceptualization of asexuality Johnson uses is close to celibacy, chastity, and virginity, but trying to avoid the religious and moral connotations; she also wants to allow asexuality to be defined in terms of preferred rather than actual sexual behavior. While Johnson dislikes the term celibacy (often putting it in quotes when she has to use it), I will call this conceptualization of asexuality preferential celibacy. Her reliance on letters to the editor for data about this group suggests the difficulty of conducting research on asexuality at that time.

3.1.5   Summary of asexuality in sexology before 2000

What this brief overview makes clear is that, while the existence of people who, in current classifications would likely be considered asexual, has been recognized in sexological literature since at least the 19th century, extremely little research had been done—some scattered case histories can be found, but prior to a paper published in 2004 (discussed below in Section 3.2), I have been unable to find any studies specifically on asexuality or lifelong HSDD, with the sole exception of Johnson's 1977 paper. While models that include asexuality have existed for some time–Storms (1980)used the term asexual and defined it in terms very similar to AVEN–only a handful of studies attempted to use this. The most plausible explanations for this state of affairs are a) a relative lack of interest in asexuality among

sex-researchers and b) a lack of availability of research participants. Whatever the reason(s), this state of affairs has changed considerably in the past 15 years.

3.2 More recent academic discourses about asexuality

After Johnson's 1977 article, the next known academic work on asexuality was a paper by Canadian psychologist Anthony Bogaert, titled "Asexuality: Prevalence and associated factors in a national probability sample," appearing in the August 2004 issue of the Journal of Sex Research.

On account of the AIDS epidemic, since the 1990s some large scale studies about sexuality have been conducted using probability sampling. One of these, the National Survey of Sexual Attitudes and Lifestyles (NATSL Wadsworth et al., 1993) was conducted in Brittan (England, Scotland, Wales) in 1990 with n>18000, making it the largest such study since Kinsey's work–and sampling methods had improved considerably in the intervening decades. One of the items asked about sexual attraction, using a 5 point bipolar scale from heterosexual to homosexual, but it also included the option, "I have never felt sexually attracted to anyone at all." In a way this is analogous to Kinsey's X: It is off the (bipolar) scale. A total of 195 participants chose this option, which was a little over 1% of those who answered this question. Using this response to operationalize asexuality, Bogaert found various correlates: as compared to non-asexuals, asexuals on average had shorter stature, greater religiosity (on one of two measures of religiosity), lower level of education, lower socioeconomic status, poorer health, asexuals were older than non-asexuals, and asexual women had a later average onset of menarche. In a theoretical article following up on this (Bogaert, 2006), he used the existence of biological correlates of asexuality as evidence for viewing it as a sexual orientation, and likewise in an empirical follow-up (Bogaert, 2013) analyzing data from NATSL-II (Erens et al., 2001), which was conducted ten years after the first NATSL. In that sample, 0.47% reported having never experienced sexual attraction. Otherwise, most of the findings of NATSAL-I were replicated, except for the results concerning adverse health, age, and

onset of menarche.[7]

When Bogaert had begun working on that paper, he had not been aware that there was a growing online asexual community (p.c.), and searching through his other publications reveals that, in fact, this paper was simply one of many analyses that he published using national probability samples with data on sexuality (e.g. Bogaert and Friesen, 2002; Bogaert, 2003, among others). Bogaert's 2004 paper, along with the emergence of an online asexual community, served as the basis for a 2004 article in New Scientist about asexuality, which proved to have an enormous impact on asexual visibility (see Chapter 4). Since publishing this article, Bogaert has produced a number of other articles on asexuality (Bogaert, 2004, 2006, 2008, 2012a, 2013) and a book (Bogaert, 2012b), and he has appeared numerous times as an expert in TV, radio, and print media. His comments about asexuality in media articles have generally been seen as favorable when discussed in online asexual discourse.

In the time since 2004, the size of the corpus of academic work on asexuality has grown enormously, and it has predominantly come from psychology, gender studies, and sociology. This can be seen in Figure 3.2, which includes all academic English language journal articles, editorials in academic journals, book chapters, monographs, MA theses, and doctoral dissertations about asexuality that I know of, through 2013. Excluded are presentations and non-English language articles. Articles posted online ahead of print are coded according to when they went online. Articles are also coded according to the academic field of the first author. Because of similarities in methodology and the small number of articles from communication (n=2), these have been grouped with gender studies. Works from English and from comparative literature have been grouped together as "literature." The category "other" includes works from linguistics (by myself, n=2) and public health (n=1).

These will be discussed much more briefly as they have not had nearly the impact on asexual organizing as Bogaert's 2004 paper had. After that paper, the earliest articles from psychologists both used mixed-methods approaches (i.e. a quantitative analysis of

---

[7]NATSAL-II had a "core" sample and an "ethnic boost" sample. For most analyses, Bogaert only reports on the "core" sample, which was only about 60% as large as the sample for NATSAL-I. Combined with the lower prevalence of asexuality, the second study generally had less statistical power. Further, the question for menarche was in a section only completed by those who previous sexual experience, further decreasing the power for asexuals.
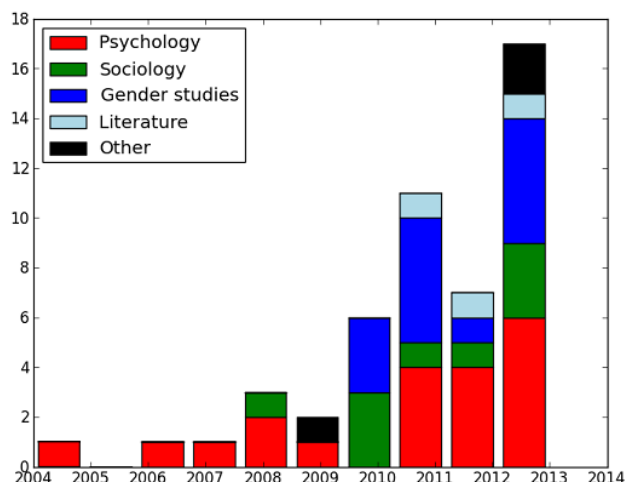
Figure 3.2: English language academic publications about asexuality, by year and discipline.

surveys, and a qualitative analysis of interviews) to engage in basic description of asexuality (Brotto et al., 2010; Prause and Graham, 2007) Other than a few studies using pre-existing large datasets (Aicken et al., 2013; Bogaert, 2004, 2013; Höglund et al., 2014; McClave, 2013; Poston and Baumle, 2010) research specifically on asexuality has recruited from online asexual communities, although some of these have recruited through other means in addition to this.

A number of people conducted qualitative research on issues pertaining to asexual identity (mostly commonly relationships and/or development of asexual identity). Some of this has resulted in journal articles and book chapters, including work by two sociologists (Carrigan, 2012; Scherrer, 2008, 2010b,a) and a group of psychologists Van Houdenhove et al. (2014). Qualitative work regarding asexual identity is also reported in a masters thesis Sundrud (2011) and a doctoral dissertation Haefner (2011).

More recently, research from psychology has been conducted on physiological aspects of asexuality Brotto and Yule (2011) possible biological correlates of asexuality Yule et al. (2014), and mental health issues pertaining to asexuality Yule et al. (2013), all comparing asexuals to other sexual orientation groups. In many cases, the methods and hypotheses were borrowed directly from research paradigms in other areas of sexual orientation research. This trend is also seen in Gazzola and Morrison's (2011) study about whether and, if so, how

frequently asexuals experience asexuality-based discrimination. (They found virtually no evidence of anti-asexual discrimination, although their sample size was small and they note that, "although many participants had disclosed their asexuality to at least one person, this aspect of their identity remains hidden from most of their proximal (e.g. family members) and distal (e.g. work peers and extended family) contacts" (p.34))

In addition to these, there has been some work from psychologists and myself discussing a number of theoretical and methodological issues (Bogaert, 2006; Brotto and Yule, 2009; Chasin, 2011; Hinderliter, 2009b), such as sampling, operational definitions, the use of questionnaires about sexual function/dysfunction, possible differences between self-identified asexuals and asexuals who have never encountered an asexual community (Chasin uses the term "potential asexuals"), and the relationship between asexuality and HSDD.

Since 2010, a sizable portion of work on asexuality has come from gender studies and communication. So far, none of the journal articles or chapters from these fields published from 2010-2013 reported any empirical research (qualitative or quantitative) on asexuality. Kim (2010, 2011) discusses issues pertaining to asexuality and disability from a disability studies perspective. Cerankowski and Milks (2010) give a general introduction to asexuality and argue that feminism and queer theory should take it seriously. Przybylo (2011) considers asexuality's potential for transforming social norms about sexuality, and Gressgård (2013, p. 180) considers the question of asexuality's potential value for queer theory, arguing that work on asexuality has the potential to "revitalise queer critique of naturalised gender and sexual identities and hetero-normativity."

Presumably, the increase in research on asexuality stems from a combination of an increase in available research participants, and from increased social interest in asexuality, both resulting from the growth on online asexual communities and some individuals in those communities trying to promote asexual visibility. The history of these communities is the subject of the next chapter.

3.3 Summary

In this chapter, I surveyed academic work about asexuality (or things that, in retrospect, we might classify as asexuality). In pre-2000 literature, I found four general ways of conceptualizing asexuality: *Asexuality as pathology*, *asexuality as a throw-away category*, *asexuality as a sexual orientation category*, and *asexuality as preferential celibacy*. However, this resulted in almost no research on the topic until after online asexual communities had started to form.

A 2004 paper by Anthony Bogaert, along with growing membership on AVEN, led to an article in *New Scientist* about asexuality, and this marked a major transition in asexuality online–a large growth in size and more media attention. Since then, there has been considerable growth in the body of academic literature about asexuality, presumably stemming from the increased social interest in asexuality along with the greater access to research participants. Primarily this work has come from three places: 1) psychology, 2) sociology, and 3) gender studies, communication, or closely related areas.

Chapter 4

A brief history of asexuality

In the previous chapter, I surveyed academic discourses regarding asexuality, demonstrating that there is research on something very similar to asexuality going back to at least the late 19th century, and that the conceptual resources necessary for creating an understanding of asexuality that is similar to what now exists in online asexual discourse has existed in Western cultures for quite some time prior to the emergence of online asexual communities. With the emergence of such communities, there has been an enormous growth in the amount of academic discourse about asexuality.

The growth of these communities and the conceptual options used in their vision of asexuality—especially in the earlier years—will be the focus of this chapter. This chapter relies primarily on archival research using newspaper archives, online archives such as the WayBack Machine, and still existing asexual websites, blogs, etc. It relies primarily on qualitative analysis of these, with a focus on "early" asexual websites. Sometimes this is supplemented with quantitative data focusing on the *where* and the *how much* of online asexual discourse. The results here largely inform the hypotheses investigated in Chapters 5– 6.

This chapter is organized as follows: In Section 4.1, I will briefly discuss mentions of asexuality in US and UK newspapers prior to socially widespread use of the internet and discuss what is known about asexuality and asexual identity prior to the formation of online asexual communities. In Section 4.2, I will examine "early" online asexual organizing when most of what existed was static content sites; in Section 4.3, I will look at growing dynamic content in online asexual communities; and in Section 4.4 I will look at how AVEN came to

dominate asexual discourse for a time and then at a subsequent increase in other locales of online asexual discourse–these three sections will primarily focus on the *where* and the *how much* of asexual discourse. In section 4.4, I briefly outline some important concepts in online asexual discourse, and I present some of my main historical hypotheses in this dissertation; I provide some qualitative evidence that these changes have occurred, and these hypotheses largely informed my choices of methodology discussed in Chapters 5 and 6.

4.1 Asexuality before the internet

Very little is known about asexual identity or organizing (if there was any) prior to the increasingly widespread use of the internet in the 2nd half of the 1990s. In Chapter 1, I argued that asexual organizing has only begun on any socially noticeable level with the growth of the internet, and gave two main lines of argument for this. First, all of the most extensive accounts of the history of asexual organizing begin with online organizing; and these include historical accounts from four different people (one of whom is myself) who had all done extensive searching for information on the topic. Second, in an online survey with 207 completed responses from asexual-identified individuals 25 or older as of January 2012, I found no evidence of pre-internet asexual organizing.

To further investigate asexuality prior to the socially widespread use of the internet, I conducted searches through the archives of *The New York Times*, *The Times* (London), and *The Guardian and The Observer* for the terms *asexual* and *asexuality* during the period of, roughly, 1950–1990, depending on the material available in the various databases. In addition to these, there is an article on the AVENwiki about pre-2000 articles based on (among other things) similar searches through the Google News Archive[1]

To transition from academic/medical discourse about asexuality to more popular dis-

---

[1]I used ProQuest Historical Newspapers to search through *The New York Times* and *The Guardian* and *The Times Digital Archive* (1785–1985) for *The Times*. The Google News Archive began adding old newspapers to their archive in 2008, but stopped adding new material in 2011 (Horn, 2011). The database, though extensive, is uneven in its coverage, and generally does not have materials from the largest newspapers. For a list of what it contains, see http://news.google.com/newspapers (Retrieved 2013-04-29). The search through The Google News archives were conducted myself and one other person, and were discussed on a listserv for people interested in the history of asexuality. This formed the basis of the article on the AVENwiki: http://www.asexuality.org/wiki/index.php?title=Pre-2000_articles, retrieved 2013-04-29.

course, it seems fitting that the two earliest popular articles I can find in which asexuality is framed in terms of sexual orientation come from a syndicated medical column. In an article about "transvestism"[2] and homosexuality, Walter Alvarez (1959b) provides anecdotal evidence from the family history of homosexual patients of his:

> Occasionally, I have found others in the family who were homosexual or what I call asexual. They were persons who seemed uninterested in either men or women, and who usually never married. For instance, a homosexual man recently said to me, 'I cannot tell you definitely that some of my brothers are homosexual; I don't know; but five of them never dated and never married.

In an article a few months later (Alvarez, 1959a), he said something similar about asexual relatives, but this time in relation to men who lose interest in sex in their 40s. In the quoted article, contrasting asexuality with homosexuality suggests it is being viewed in terms of sexual orientation. In both articles, it appears to be viewed as pathology—not as a pathology of principle interest, but rather of interest because it is taken as evidence for the congenital basis of other pathologies[3] Also in both articles, asexuality is defined in terms of lack of interest in either men or women, with never marrying being a common feature.

Examining the use of *asexual* and *asexuality* in the newspaper databases, there are some instances in both US and UK newspapers where *asexual* is used and the context suggests it is being used as a sexual orientation term. Such uses are infrequent and—other than the two Alvarez articles—generally begin to be found in the late 1960s to early 1970s. For instance, the *asexual* is used in a list of sexual orientation in a comical article describing a fake computer dating service (Guardian, 1969), in a review of *Jesus Christ Superstar* (Boyd, 1971) and in an article quoting President Carter's press secretary (Smart, 1977). There are also cases where *asexual* is used in contrast with *homosexual.* Someone claimed that Lawrence of Arabia was not homosexual but asexual (Moorhouse, 1969). In a review of a staged version of *Ballad of the Sad Café* one character was described as asexual and

---

[2] The group would today be called transsexuals.

[3] In the case of homosexuality, he considers asexual relatives relevant on account of his view that homosexuality "is commonly inherited as a variant of psychosis."

another as homosexual (Billington, 1971) A serial killer, denying that his homicides were sexually motivated, said that "he had some homosexual relationships but believed he was asexual" (Davies, 1983, p.1) Aside from articles in advice columns, I know of only three newspaper articles primarily about asexuality during this period. The first two both appear in *The Village Voice* (Flaherty, 1971; Bell, 1978) The first (Flaherty, 1971) and third (Hoppe, 1979), are satire while the second (Bell, 1978) is a meandering piece that speculates whether Ed Koch, then mayor of New York, might be asexual and muses whether this is indicative of a general move towards asexuality with many people being burnt out from the ubiquity of sex.

The place where asexuality appears to be most often mentioned in US newspapers prior to the emergence of online asexual communities is advice columns. For instance, Ann (Landers, 1979, p.21) said to a letter writer, "You are probably 'asexual' which means your sex drive is so low it is virtually non-existent." A column in Dear Abby (Van Buren, 1981, p.7) included a letter from someone identifying as asexual, who interpreted it as a sexual orientation category: "You have been, over the years, a great friend of hetero-, homo-, and even bisexuals. How about a good word for those of us who are genuinely asexual? We are truly a forgotten minority." A few similar examples can be found in the 1970s, 1980s, and 1990s, but not many—the AVENwiki only identifies nine such articles. These demonstrate that the basic conceptual building blocks for understanding asexuality as a sexual orientation were available in the general population in the US and UK for several decades before the emergence of online asexual communities, and they also provide evidence that there were at least some people identifying as asexual during that time. However, they provide no evidence of asexual communities or organizations (the two satirical articles invented fictitious asexual organizations). And while such articles likely led some people to believe that asexuals exist, and may have led some people to consider themselves asexual, they provided no possibility for asexuals to communicate directly with other asexuals. Socially widespread use of the internet has greatly changed this situation.

4.2 Early internet asexuality

In doing research on the early history of asexuality on the internet, various sources of information are possible. In some cases, websites are still active and have old posts within their archives. In other cases, the websites are no longer active, but are archived on archive.org's WayBack Machine, which has "over 150 billion web pages archived from 1996 to a few months ago,"[4]although they generally do not archive forum data. Its archiving of webpages uses an opt-out policy, so many early asexual sites are archived there, Other sources of information about early online asexual history include a number of pages on the AVENwiki, a college term paper and a podcast by Jay (2003, 2006), the founder of asexuality.org, and a presentation by Titman (2012), the founder of the asexuality community on LiveJournal. In a few instances, I have contacted people via email.

Little is known about the earliest online organizing of asexuals. In a survey I conducted (see Chapter 6), participants were asked when they first joined an online asexual community, and one reported finding some other asexuals in an unrelated usenet group in 1995. According to the AVENwiki's page about the history of asexuality,

> In the early and mid 90s, people would occasionally post on sexuality-related newsgroups expressing their identification with asexuality or their lack of sexual attraction, but search functions were not effective enough, making it difficult to locate people with similar experiences.[5]

For evidence of this, three Usenet threads are cited, along with a paper Jay (2003) wrote in college that he posted on AVEN and is now archived on the WayBack Machine.[6] In March 1998, Regena English published an e-book called *Leather Spinsters and their Degrees of Asexuality* based on interviews with and letters by a rather diverse group of women that English calls Leather Spinsters, a term used to try to reclaim spinster and give it a positive connotation. English's leather spinster category combines women who prefer to be single and

---

[4]Retrieved 2013-01-14 from http://archive.org/web/web.php
[5]Retrieved 2012-01-04 from http://asexuality.org/wiki/index.php?title=History
[6]Many posts from Usenet are currently archived in google groups. For more information see http://en.wikipedia.org/wiki/Usenet.

women who do "not identify with heterosexuality, homosexuality, or bisexuality"(English, 1998, p.4). The interviews and letters suggest that spinster (as a positive term) and asexual were sometimes terms that people used on their own, and were sometimes terms that seemed new to respondents. In the introduction, she explains that the book was written "to satisfy an urge on my part to meet others like myself, happiest when alone...I wondered if I was the only one that felt that way" (English, 1998, p.4). She first attempted to do research through the library:

> I read all I could on lesbianism and bisexuality on the hopes of analyzing and applying their struggles to make sense out of my own. It was during one of my many discussions with my sister that I actually gave a name to perpetually single women, spinsters. I heard that word all my life but I had never heard it as a positive term to praise and uplift any unmarried woman. I went back to the library to research spinsters, and to my amazement found that such books were almost non-existent" (English, 1998, p.4).

Thus, she sought to interview women for her book. To do this, she used classified ads in newspapers, searched the internet, and spent considerable time in internet chatrooms. After publishing her e-book, she created a website (leatherspinsters.com) and in August 1998, the site began hosting the Leather Spinsters Newsletter. Sometime later, a forum (now defunct) for single women was created to go with the newsletter.

English's book and website appear to have had relatively little impact on subsequent asexual discourse, but the great difficulty that English had in finding existing published information on asexuality is illustrative of the lack of information that was available until recently, and the extreme difficulty that was involved in finding asexuals at the time—and the use of the internet for finding many of her interviewees further demonstrates the importance of the internet in being able to find many asexuals.

One of the first online articles about asexuality was called "My life as an Amoeba" and appeared in an e-zine called StarNet Dispatches.[7] The use of amoeba is a pun based

[7]Retrieved 2012-06-19 from

on the asexual reproduction meaning of asexual. Jay (2003) and Titman (2012) suggest that this connection was especially salient to people at the time who tried to learn about asexuality/find other asexuals by entering 'asexual' into search engines and mostly got pages about asexual reproduction. Like many online pieces, this article had a comments page.[8] Perhaps different from many online articles of the time, this one continued to get comments for years after first being published, often by asexuals or by people wondering if they are asexual or not. Likely this was because of the difficulty of finding places to talk about these matters.

Titman also suggests that a major factor necessary for the growth of online asexual organizing was improved search functionality, citing as evidence the fact that, although there were numerous comments made before the last time the article was archived (2003), it was about two years before many more comments started appearing other than those made soon after the article first appeared. An alternate hypothesis is that, at that time, the number of internet users was dramatically increasing. Limits in available data make it difficult to test between these, and it is very possible that both are correct.

In the late 1990s and early 2000s, a few more "asexual" sites and/or articles popped up around the internet, although there had not been a convergence on the word asexual. A website called The Organization For Antisexualism (OFA) was created in April 1999 and was hosted on Geocities. At some point, OFA developed a listserv, but the content of this is not in the public domain. According to a moderator on that listserv (Swankivy, p.c.), "There was very little discussion and it felt mostly like just an acknowledgment of other people existing who weren't interested in sex." Swankivy cites that the reasons for it closing down were the increasing popularity of AVEN, along with the founder of OFA not being able to devote much time to it due to personal reasons.

In 1998, Swankivy posted on an article called "The Nonsexuality Rant" on her AOL website. Swankivy has since then become a relatively well-known person in asexual communities,

---

http://web.archive.org/web/20020601161040/http://dispatches.azstarnet.com/zoe/amoeba.htm

[8]Retrieved 2012-06-19 from

http://web.archive.org/web/19980114143639/http://dispatches.azstarnet.com/zoe/amoeba2.htm

and she has periodically updated the article, with the most recent version titled "Asexuality Top 10."[9] It was because of this article that Swankivy was contacted by the creator of OFA and eventually became a moderator for their listserv. Another "early" article was called "The Fourth Sexuality" and is dated November 13, 2000.[10] Some of these early asexual locales linked to each other (and people corresponded privately), thus creating more—though still very limited—possibilities for asexuals to communicate with each other.[11]

In 2001 and 2002, a few more static content asexual sites were created. In spring 2001, while a freshman in college, David Jay had created a static content site on his university webspace, called the Asexual Visibility and Education Network (AVEN). It had some of his own thoughts about the topic as well as contact information. One person—who I will refer to as TG—had a few pages about asexualism on her Geocities site.[12] She added these in July 2001, but they appear to have been removed sometime in 2002.

Another person made a Geocities site called "The Asexual Manifesto" which had a few pieces the author had written. This person also obtained the domain asexualmanifesto.org, which had a link to the geocities site. The Geocities site remained online until Geocities was closed, but asexualmanifesto.org only appears to have been used around 2001–2003.[13]

Sometime in 2002, asexuality.net was established as a static content site about asexuality. Most likely, the site was launched in February 2002, and although it was probably the best designed asexuality site (with the best domain name) at the time, available evidence suggests is lasted for less than a year.[14] Many of these sites linked to each other, and their development

---

[9]Retrieved 2012-06-20 from http://swankivy.com/writing/essays/philosophy/asexual.html.

[10]http://web.archive.org/web/20020808195605/http://oak.cats.ohiou.edu/lb122098/fourthsexuality.html

[11]Evidence for this comes from a paragraph added to the Nonsexuality Rant sometime in early 2001: "Here are some connections: I belong to the Organization for Antisexualism (OFA) [link] and the Celibate Webring [link], and I have received mail from others like myself. Here is an essay sent to me by a girl named Laura about our shared sexual orientation of nada: The Fourth Sexuality [link]. Please read it and tell us what you think. And here is a cartoon [link] my friend Kari made commenting on our celibate nature; tell us what you think! We are not alone and we are not crazy!" Retrieved 2012-01-04 from http://web.archive.org/web/20010215104910/http://mypage.goplay.com/venusgiri/nonsex.html

[12]The only contact information I had was a 12 year old email address. I tried to contact her, but, unsurprisingly, was not able to, so I will use this name for her online pseudonym.

[13]The main evidence is dates of when it was archived on the WayBack Machine. Murphy et al. (2007) provide evidence of the validity of using the WayBack Machine for dating. However, it is not very precise because there may be a delay of several months between when a change was made and when site was next archived.

[14]Available information comes from the WayBack Machine and posts on Haven for the Human Amoeba. The domain name had been obtained prior to Jan 12, 2002, and the first content was posted sometime between that date and February 19th. The Way Back Machine's last record of it existing is September 26th of that year, and by February 8th of the next year it

coincided with the growth of asexual dynamic content.

4.3   Increasing asexual dynamic content

Probably the oldest online asexual community with still existent records is a Yahoo! Group called Haven For the Human Amoeba, created in October 2000 and listed in the category Romance & Relationships > Relationships > Abstinence. The group's name is probably a reference to "My life as an Amoeba" which the group's creator linked to in an early post (Titman, 2012). Initially the message archive was completely public, but sometime in 2006 or 2007, this was changed so that, to read posts, it is necessary to join the group, although all that is needed to join is a Yahoo! ID (no confirmation from anyone in the group is necessary). Non-members can view a description of the group and information on activity (i.e. how many posts per month). For this reason, quotes will only be used in cases where I have permission from the author. In addition to its email list, there was a chat room— this data has not been preserved, but sometimes reference to it is made in the posts. To illustrate the role this site has played in online asexual discourse, the number of posts per month through 2012 is given in Figure 4.1.[15]

For about the first eight months of its existence, posting was infrequent. Sometimes the owner would post something about an article or a book of relevance or a new member would give a self-introduction. Sometimes a few people would respond, and then there would be no discussion for some time. In July 2001, the group reached a critical mass and (still extant) online asexual discourse has been more-or-less continuously created ever since, although not always in this Yahoo! Group. Haven for the Human Amoeba was most active in the year or so after that critical mass was reached. After that, there were varying periods of activity and inactivity, and then extremely little since 2007, with the exception of one month. Online asexual discourse had largely moved elsewhere.

was gone. Most likely, the owner had purchased the domain and hosting for one year and chose not to renew it. My guess is that, when created, the site filled an unmet need for a well-designed asexual site with a good domain name, but that in its first year, AVEN was moved to asexuality.org, and its static content and forums made the owner feel asexuality.net was no longer necessary.

[15]Retrieved 2012-06-20 from http://groups.yahoo.com/group/havenforthehumanamoeba/
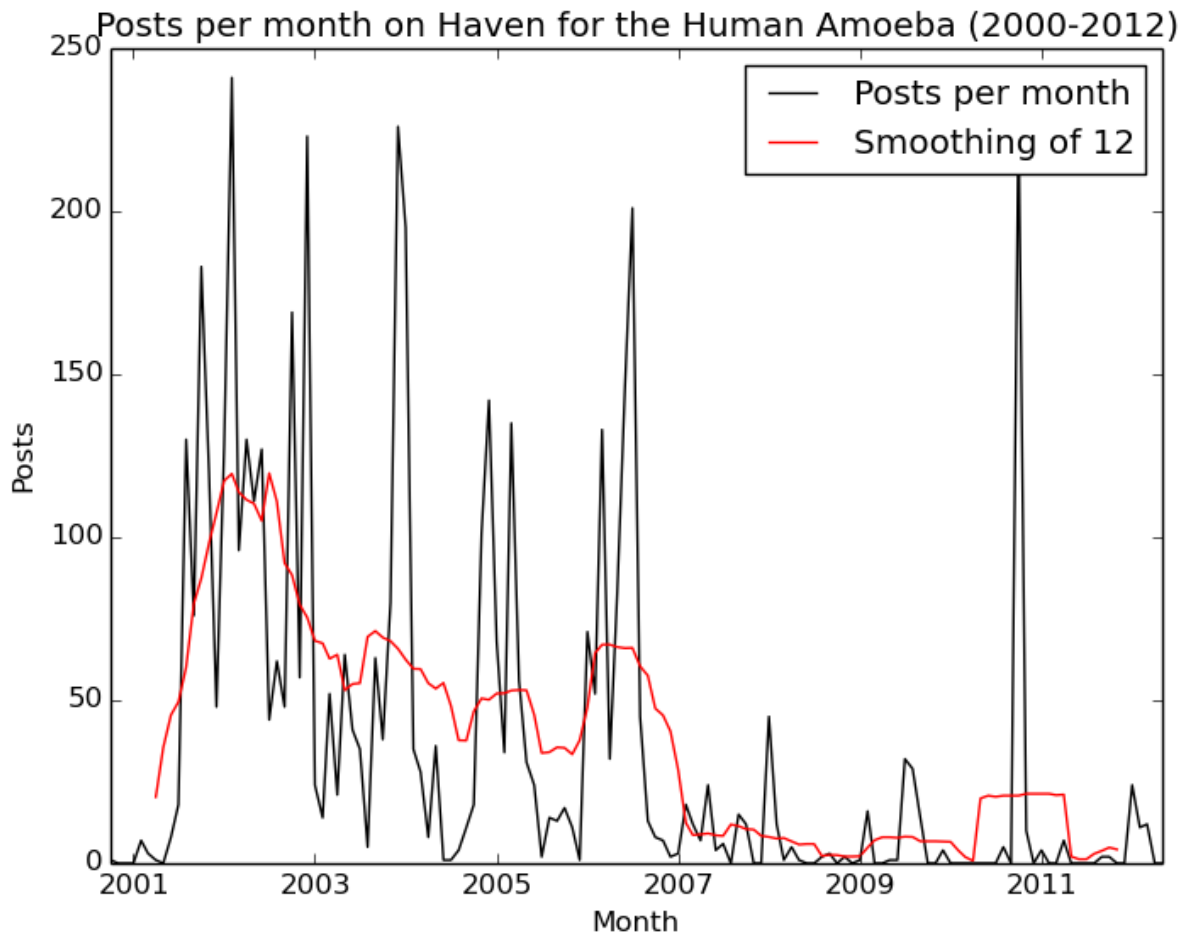
Figure 4.1: Number of messages per month on Haven for the Human Amoeba.

The growth of asexual discourse on Haven for the Human Amoeba in 2001–2002 had important connections with the static content sites discussed in Section 4.2 above. For a time, asexualmanifesto.org was linked to in Haven for the Human Amoeba's "about" section, and appears to have been written by the group's owner. TG added a page about asexuality to her site a few weeks after joining Haven for the Human Amoeba in July 2001, and it was a few weeks later when she started the asexual webring and contacted David Jay about his site AVEN.

In spring 2002, two new asexual locales arose independently of each other, and these later became, for several years, the two largest online asexual locales: AVEN's forums and a LiveJournal (LJ) community called Asexuality (asexuality.livejournal.com). The LJ community was created by Nat Titman to be "an explicitly sex-positive asexuality discussion community,[16] in response to another LJ community, called Asexuals, which Titman saw as made up mostly of celibate people with negative views of sex.

On LiveJournal, each member has their[17] own page, on which they can keep a journal/diary/blog. They also have a list of "friends" (others on LiveJournal) and they can join "communities" which generally focus on some common area of interest. Individuals can post to their journal, they can post to a community, and they can comment on other people's posts and also comment on other comments. For the Asexuality community, while it is possible to join LJ specifically to participate in this community, many participants already had accounts on LJ and can participate in the Asexuality community (or simply follow it) with accounts that they already had and also used for other things.

Figure 4.2 displays the number of entries per month on LiveJournal Asexuality from 2002 through May 2013.[18] As can be seen, there was a relatively low level of posting for the first few years, and then the community was most active in 2005 and 2006. Since then, activity

---

[16]Retrieved 2011-02-26 from http://asexuality.org/wiki/index.php?title=Livejournal_Asexuality.

[17]A relatively large portion of asexuals—including Titman—have a gender identity that is neither male nor female, and often prefer the pronouns they/them/their. As such, the academically preferred he/she seem in appropriate.

[18]The data are based on a web-crawl conducted in June 2013. Prior to learning how to do this, on 2012-06-20, I did manual data entry using the *archive* page and pages linked to in it from http://asexuality.livejournal.com/calendar. The numbers differed in some cases, with the manual input always being larger when there was a disagreement. Most likely this is due to people deleting accounts and content between these two times, thus suggesting that the figures shown somewhat underestimate things. While the figure excludes comments, entries per month are correlated with comments per month at approximately r=0.944, and with words per month (in comments plus entries) at about r=0.92
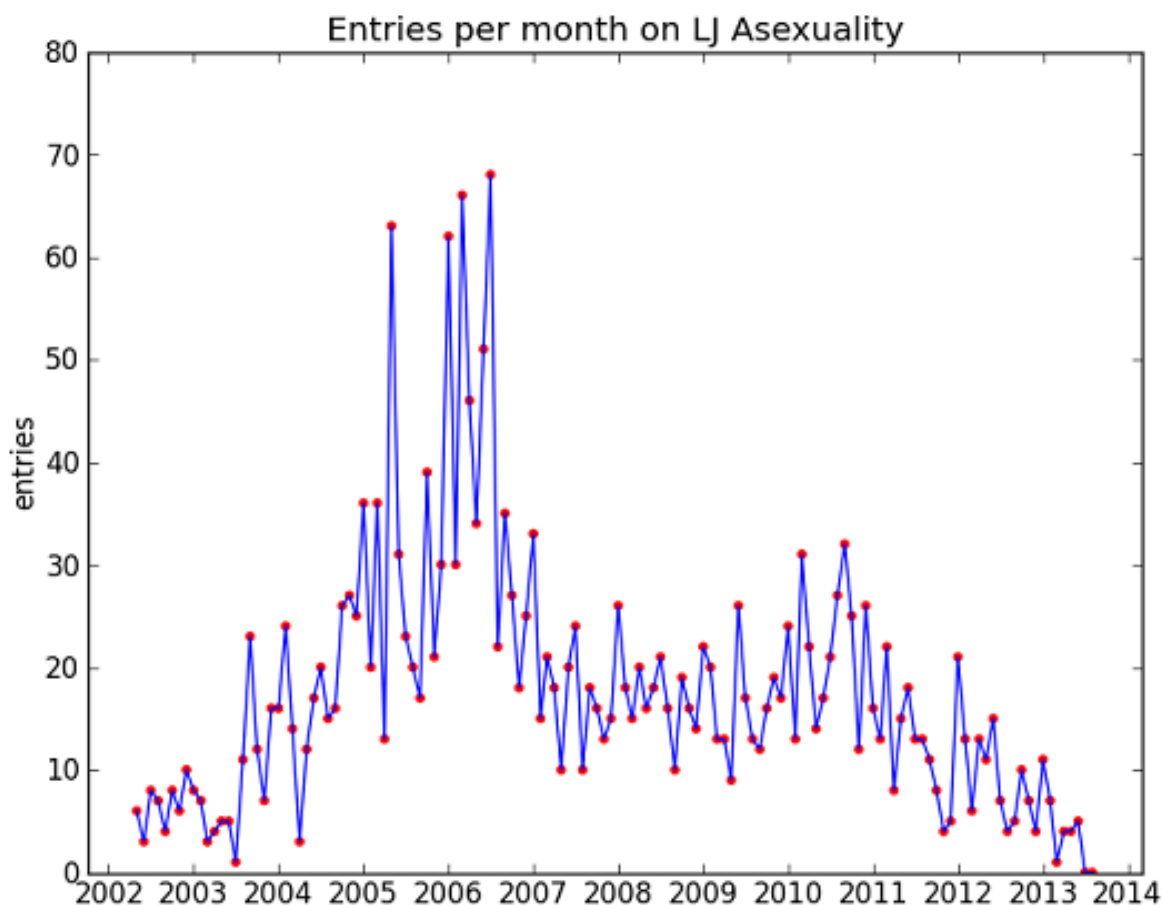
Figure 4.2: Number of entries per month on LiveJournal Asexuality. Red dots indicate data points.

continued for a few years, but at a lower rate. The most recent figures indicate only a few posts are made per month. This trend is probably common among many websites, but it does contrast with the trend on AVEN, which had a surge in usage in late 2004, but has not experienced a subsequent decline.

After David Jay joined Haven for the Human Amoeba, he actively posted there. In late December 2001, he created a listserv for people interested in asexual activism, feeling that Haven for the Human Amoeba discussion focused more on chatting than on organizing. About two weeks later, in early January, he announced that he was in the process of acquiring the domain asexuality.org and was planning to move AVEN there. Jay (2003) explains that in Haven for the Human Amoeba, every member got every e-mail sent to the group and so sustained communication about asexuality there tended to flood people's inboxes.

Further, there was no way to have multiple threads going at once. As such, it was a very limiting way to communicate about asexuality, especially given a desire to get new members. David Jay finally acquired asexuality.org in late May, making it the new home of AVEN and setting up forums on May 29th, 2002, about a month after the creation of the LJ asexuality community.

Figure 4.3 shows the number of posts per month for most of AVEN's forums viewable to non-members.[19] The general trend has been an increase in content over time.

To understand some particular aspects of this growth, it is important to examine the impact of media pieces on asexual visibility. There have been a large number of media pieces—most commonly text (print and/or online), but also radio and television. If some media piece results in a great deal of visibility, then we would expect a large influx of new users directly following its publication. Thus, to see the role of media Figure 4.4 shows the number of new users per day in my AVEN data.[20]

As can be seen in Figure 4.3, the number of posts per month generally increased slowly from 2002 until October 2004, when there was a large increase. The sharp decline in early 2010 is due to a data loss of about 20 days, and the sharp rise in posts in the latter July – November 2010 (which does not obviously correspond to an increase in new members), is largely due to one highly anomalous thread.[21] Examining the number of new users per day, there are three very obvious spikes: one in October 2004, one in March 2006, and one in January 2012. All three correspond with major media pieces.

The first peak is on October 15 2004 (n=56), although the rise began the day before (n=24), corresponding with a feature article on asexuality appeared in New Scientist on October 14(Westphal, 2004). Along with this article, CNN published an article on October 14 (CNN, 2004) and The Guardian did on October 13 (Radford, 2004), both citing the article in New Scientist. In asexual discourse, this seems to be seen as the first really major media attention asexuality had received. Comparing this with the posts per month, we see that,

---

[19]This is for the part of the AVEN forums I am labeling as the "core". See Section 5.1 for details.

[20]This is not equivalent to new accounts per day, because many accounts never post. Because of the way my data was collected, I only have information for accounts that posted at least one word in my corpus.

[21]Interested readers can find the evidence for this in 4.3.
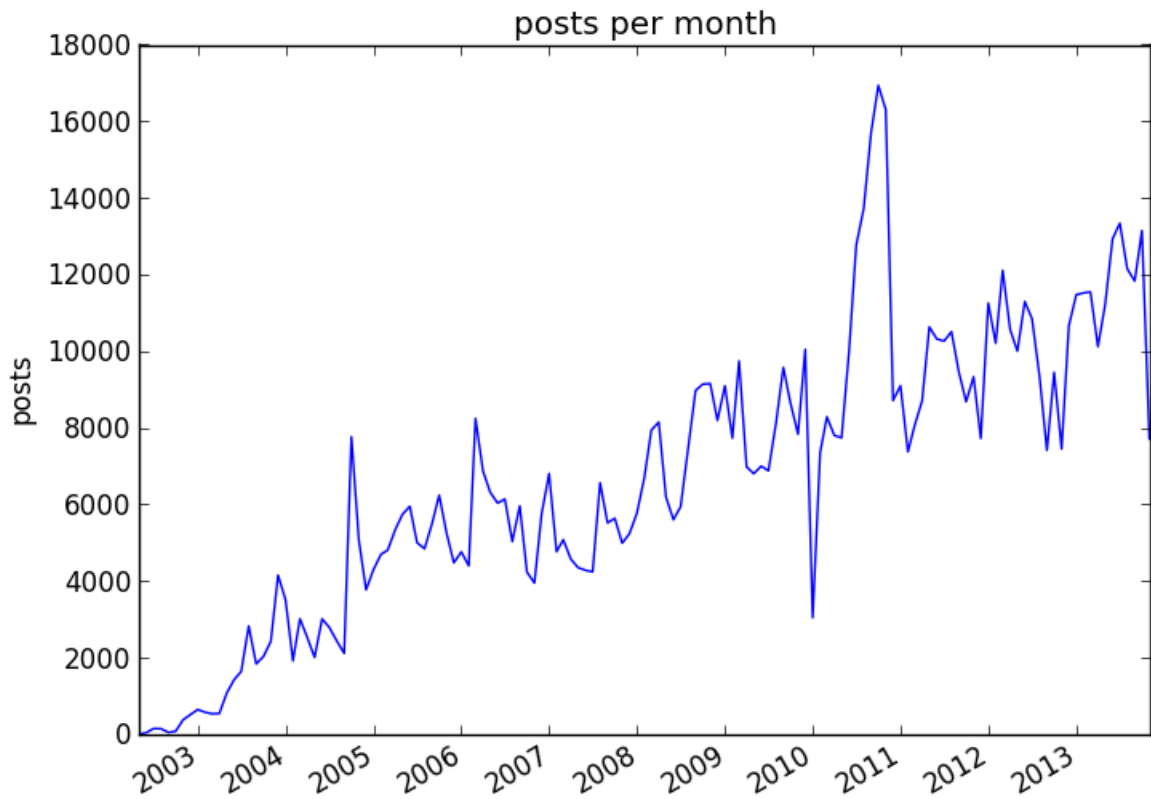
Figure 4.3: Posts per day in most of AVEN's public forums
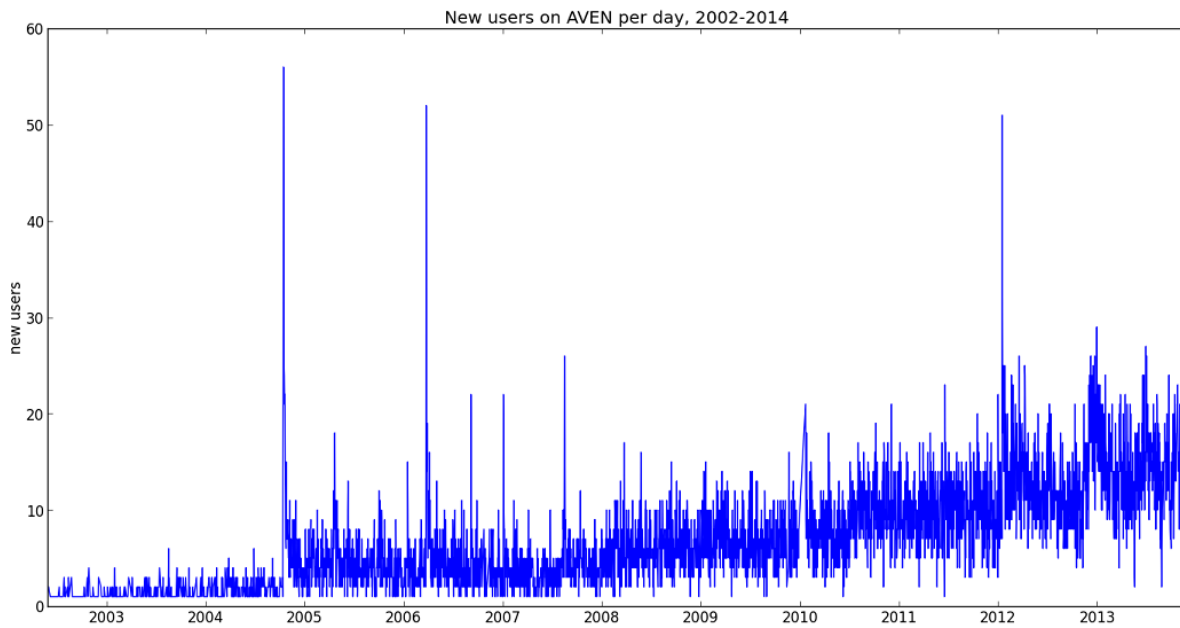
Figure 4.4: New users per day in most of AVEN's forums

50

after things had settled down following the initial surge, the number of posts per day was about twice what it had been beforehand.

The second peak is on March 25 2006, corresponding to a segment on asexuality on the news program 20/20 on March 24. The third major peak a segment about asexuality on a BBC program called "How Sex Works" (January 16) and a corresponding BBC News Magazine article (January 17 Wallis, 2012). The other smaller spikes in 2006-2007 also correspond to media attention: The spike on September 6, 2006 reflects a rerun of the 20/20 episode. The one on January 4, 2007 corresponds to an episode about asexuality on the Montel Williams Show, and the one on August 17, 2007 coincides with a rerun of that episode. However, none of these was an obvious turning point the way that the New Scientist article was.

What made asexuality newsworthy for New Scientist, a science journalism publication, was the combination of two things: the publication of Bogaert's 2004 paper in the August 2004 issue of the *Journal of Sex Research* and a growing online community of people identifying as asexual. Although Bogaert reported a number of correlates of asexuality (discussed in Chapter 3), media articles often ignore most of these and as well as Bogaert's very good discussion of the limitations of his study. They often latch on to one thing: the 1% figure for the prevalence of asexuality.[22] The article also quoted Nicole Prause, whose work was mentioned in Chapter 3, who used her data to suggest that asexuality is not pathological.

4.4 The dominance of AVEN and subsequent growth of other online asexual spaces

To show an approximate distribution over time of where online asexual discourse was occurring, Figure 4.5 shows the number of words per month in Haven for the Human Amoeba, LJ Asexuality, and AVEN's main forums from July 2001 through June 2003. (See Chapter 5.1

---

[22]Between the above mentioned articles in New Scientist, CNN, The Guardian, two mention the 1% figure—it's in the title of the CNN article—but none mention any of these other findings. My impression from reading similar articles over the years is that this is fairly typical.
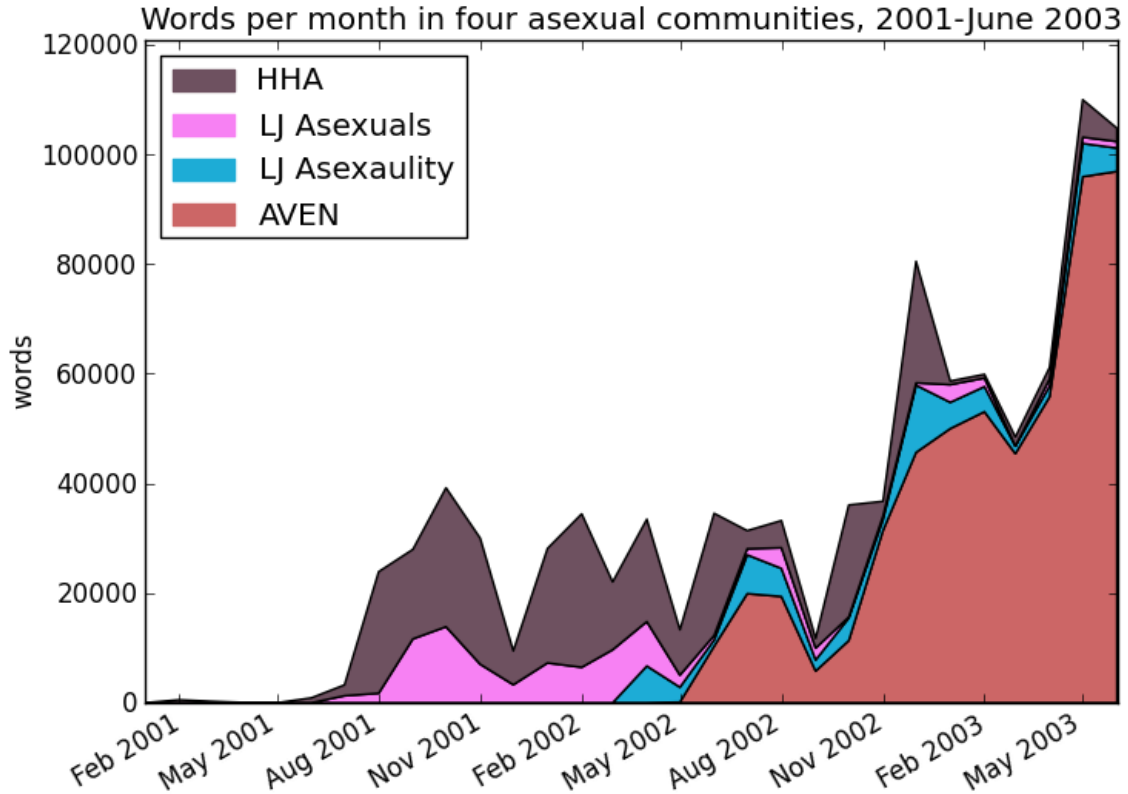
Figure 4.5: Words per month in four "early" online asexual communities.

for details about these corpora.) While the creation of the AVEN forums did not suddenly cause other English language asexual sites to vanish, it can be seen that the AVEN forums came to dominate by early 2003.

At the end of May 2005, AVEN had a major software upgrade for both the forums and the static content, and the extent of AVEN's dominance among online asexual communities can see be through a change in the front page, and the (lack of) response in the forums. Previously, AVEN's front page had a link to a links page, which linked to a number of the sites and articles mentioned above and plus a few others. Along with the 2005 upgrade, much of the introductory material was rewritten, and there was no longer a links page. When the upgrade was announced, a thread was started in which members were asked if anything was missing that they would like to have re-added. No one mentioned the links page. Whether the change was intentional or simply an oversight, no one seemed to miss it. After this,

AVEN did not have a links page until early 2009, following another re-design of the site.

Starting in 2004–2006, AVEN began to have an increasing amount of non-forum content. In October 2004, the "AVENblog" was started, and it hosted articles by various people on issues pertaining to asexuality. It was later renamed Asexual Perspectives, and since 2005, it has been featured on the front page, although it is now rarely updated. In September 2006, AVEN added a wiki. Because of its *Lexicon* page,[23] as well as other pages intended as reference articles, it is likely that creating and updating the wiki has played a role in asexuality-related lexical developments, although this is difficult to measure.

In addition the wiki is an especially useful source of information about the development of asexual terminology for two reasons. First wikis have a "history" tab, that enables users to look at a snapshot of that page each time it was edited, along with information about who, what, and when. Secondly, the goal of the wiki is to be useful reference, reflecting what seem to be consensus opinions (at least at the time it was created).

Also beginning in September 2006, AVEN began producing an online newsletter, called AVENues. It was initially monthly and then (roughly) bi-monthly through March 2009, and then intermittently since October 2010. Because of its "news" section and other features, it is a good source of historical information.

In early 2008, a new asexual forum called Apositive (apositive.org) was created with the goal of fostering discussion of issues pertaining to asexuality in a context where it could be assumed that everyone involved already had general familiarity with asexuality. The people who created it and many of its initial members were then-present and past AVEN admods (a hypernym for administrators and moderators on AVEN). It was quite active for the first several months, but since then activity has been slower. Also in 2008, a number of new asexual blogs were started (Other than the "AVEN blog", I know of only three asexual blogs from before 2008, only one of which was still active in 2008). Combined, these gave rise to a sense of increasing online asexual discussion off of AVEN, as illustrated by a change to AVENues in the August 2008 issue. A "letter from the editor" explains, with some degree

---

[23]Retrieved 2013-05-09 from http://asexuality.org/wiki/index.php?title=Lexicon

of exaggeration:

> Once upon a time, AVEN was the only place for lonely and confused asexuals to go to learn about others like them. That time is over. By my last count, besides AVEN there are four other forums for asexuals, more than a dozen asexual blogs, a large number of asexual communities on LiveJournal, Facebook, and similar social networking sites, and six websites devoted to personals and dating that cater specifically to asexuals, as well as several other informative websites. That's just the count in the English language (Hallucigenia, 2008, p.2).

This impression is confirmed by looking at the number of words per month in my Apositive corpus and asexual blog corpus (see Chapter 5 for details), shown in Figure 4.6. The decline of Apositive seems to coincide with the sudden growth in asexual blogs, and the relationship may be a causal one, especially because both seem to have filled a similar niche. However, the number of combined words per month in my data for these never exceeds 16% of the words that month in my "core" AVEN data.[24]

In 2010, a YouTube collaboration channel called Hot Pieces of Ace was created in which, each week, a number of regular vloggers would post a video about a specific topic for that week. Over the next year and a half, they posted about 450 videos. In 2011 another YouTube collaboration channel called The Dapper Ace was created using the same model, producing 171 videos before it finished.

While there have been a few other asexual forums, none has been able to have much long term success, with the exception of Ace-Book, an asexuality personals/social networking site, but most of its content can only be seen by members. For a number of years, AVEN was felt to be the hub of "the asexual community," although in the past few years, this perception has changed, largely through an increased asexual presence on tumblr.com, which "is a microblogging platform and social networking website. [25] In March 2014, Alexa ranked Tumblr as the 35th most visited website in the world, and the 18th most visited in the United

---

[24]I am excluding January 2010—all changes to the AVEN databases for over half of the month were lost in a data-loss.
[25]Retrieved 2012-08-03 from http://en.wikipedia.org/wiki/Tumblr
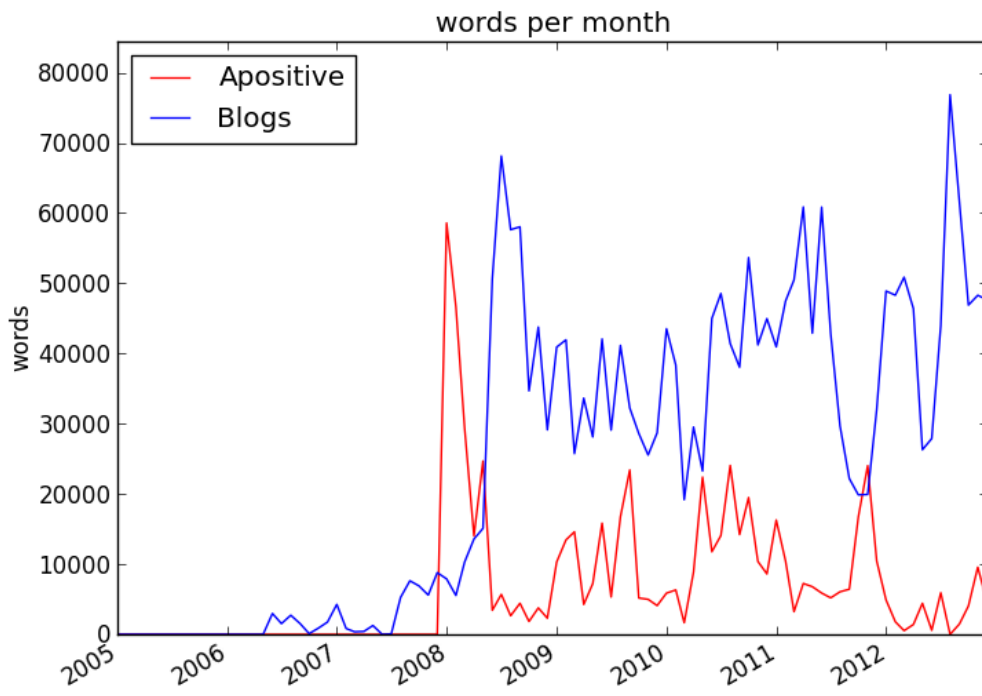
Figure 4.6: Words per month in the Apositive and asexual blog data.

Place	Percent
Tumblr	45.7
AVEN forums	34.5
Asexual friends	27.5
Livejournal asexuality community	12.0
YouTube collaboration channels	10.1
Blogs	9.9
Other forums (besides AVEN)	6.2
AVEN chat	6.2
LGBT college/university campus organizations:	4.6
In person meet-ups	4.5
Asexual & asexual-friendly dating sites:	2.8
Other text, voice, or video chat (besides AVEN chat)	1.8

Table 4.1: Usage of online asexual spaces in 2011, based on Miller (2012)

States.

Users can create their own blog (often called their tumblr) just as with other blogging platforms, but different from many, Tumblr allows people to reblog others' posts. Further, rather than using traditional comment boxes, commenting on another tumblr causes the commented on post along, with the comment, to appear on one's own tumblr. For various reasons, it is often especially difficult to find old content on tumblr. Further, individuals who post about asexuality on their tumblr often have asexuality-related-content as a minority of content on their tumblr, which makes it extremely difficult to count how many people are involved in the production of asexual discourse on tumblr, to count how much discourse is being produced, or to create a corpus from it. Besides anecdotal evidence, Miller (2012) provides relevant quantitative data. His data come from a 15 question survey run by some members of various asexual communities to get some politically useful numbers and that was promoted in as many asexual spaces as possible, resulting in 3430 usable responses. One question asked "Where do you connect with the asexual community?" The results are listed in order in Table 4.1. The fact that the most commonly chosen response is Tumblr provides quantitative evidence confirming the impression that the micro-blogging platform has become a main place for online asexual discourse.[26]

---

[26]The issue of operationalizing "asexual discourse" on Tumblr is also much more difficult than on AVEN or the LJ asexuality community, where a very simple operationalization of asexual discourse is possible: Discourse that occurs in an explicitly asexual space is asexual discourse. I discuss the theoretical significance of this several times below.

4.5 Overview of asexual concepts and terms

So far in this chapter, I have given a broad overview of the when, the where, and the how much of online asexual discourse in order to contextualize later discussion of changes in asexual discourse over time and to demonstrate the crucial role that the internet has played in the creation of asexual communities and discourse. In this section, I intend to give an overview of the what, specifically with terms and concepts that are important for understanding asexual discourse. One of the major hypotheses for my dissertation is that there have been two broad shifts in online asexual discourse, as well as some smaller shifts. The first of the broad shifts I will call "the AVEN shift" and the second "the rise of intermediate categories." In this section, I will give an overview of these hypotheses and qualitative data that supports them. In addition, I will discuss two other groups of terms: romantic orientation terminology and the terms *repulsed* and *indifferent*

4.5.1 The AVEN shift

This shift involves, to some extent, a standardization of various aspects of conceptualizing asexuality, corresponding to how AVEN's introductory material presented asexuality. This was not an entirely top-down process, as it also involves changes over time in how AVEN has presented asexuality.

In its introductory material, asexuality.org has always prominently displayed a definition of asexuality, defining it in terms of not experiencing sexual attraction, and it has made a clear distinction between asexuality and celibacy. The various aspects of what I am calling the AVEN shift have consistently appeared in AVEN's introductory material since at least mid-2002, although how prominent they have been has varied, and some parts of AVEN's introductory material was modified on account of it not being widely accepted by members of the AVEN forums.

In the AVEN shift, there was a shift in four main issues. First, asexuality generally came to be defined in terms of not experiencing sexual attraction and is seen in terms of sexual

orientation; combined with this is a complex relationship with LGBT groups as possible allies. Second, asexuality is distanced from/contrasted with celibacy. Third, the term *asexual* came to be the dominant term for describing members of this group, with alternatives (i.e. *abstinent*, *celibate*, *nonsexual*) decreasing in popularity. Fourth, in presentations of asexuality, two main kinds of diversity among asexuals are highlighted.

The first of these involves interest in romantic relationships-it is possible to be asexual and be interested in (a) romantic relationship(s). Second, being asexual is consistent with masturbation–sometimes a concept of undirected sex-drive (or something similar) is proposed to explain this. During the course of AVEN's existence, there have been some changes and developments in terminology for expressing romantic interest, but there has consistently been some sort of terminology for expressing what is now widely called "romantic orientation." (See Section 4.5.3 for details.) Based on these two ideas, in the current terminology, it is held that some asexuals can be *heteroromantic*, *homoromantic*, *biromantic*, or *aromantic* (more recently, *panromantic* has been added to this list). Further, there are asexuals with a sex-drive and asexuals without a sex-drive, although the terminology for this contrast appears to be more variable and less frequently used.

The first line of evidence for such a shift is that some of the earliest online articles (discussed above) explicitly mentioned the lack of standard terminology. The OFA's website noted that "there are a various number of terms that can be considered what we are," mentioning several: antisexual (which that author preferred), asexual (which the author disliked because of ambiguity with lacking sexual organs), celibacy/celibate (which the author used most often in explaining antisexualism, but felt that a problem was that it is "very broad, since it doesn't say why or how the person is celibate"), and abstinence (which the author objected to because it "can also be used to describe abstaining from anything.")[27] The lack of standard terminology, and the various lexical options, were also discussed in "The fourth sexuality":

---

[27]Retrieved 2012-02-23 from http://web.archive.org/web/20020808195605/http://oak.cats.ohiou.edu/lb122098/fourthsexuality.html

> Nonsexual, antisexual, asexual, celibate–these are all terms for what I am, though they have different connotations. Antisexual suggests that I think no one should have sex, which is not the case. If you're prepared to suffer the consequences, who am I to stop you? Asexual makes one sound like an amoeba, but my gynecologist says I am most definitely female. Celibate has religious connotations, suggesting that I'm making some kind of conscious effort to stay "pure." Nope. I'm making no effort; this is just how I am. I prefer nonsexual; I'm just not interested.[28]

These discussions also indicate that asexuality was positioned as being close to celibacy (that is, not having sex). In itself, this should not be surprising, as a number of sexual orientation/sexual identity terms are polysemous between behavior and patterns of attraction/desire (i.e. homosexuality, bisexuality, heterosexuality, pedophilia, fetishism etc.) However, it contrasts with what (impressionistically) is current practice: To emphasize the difference between asexuality and celibacy in presentations of asexuality.

A second line of evidence for the AVEN shift is from changes made to "The nonsexuality rant" and the introduction to Haven for the Human Amoeba. In March 2005, Haven for the Human Amoeba's description was the same as when the site was archived in December 2002: "Welcome to Yahoo's only club for asexual people and their friends. Thanks to everyone who has posted for keeping this place alive and informative."30

> By June 2005, the description had been rewritten: So what exactly is asexuality? It's a lack of sexual attraction. Asexuals are generally very different from one another: some experience romantic attraction, some don't. Some experience arousal, some don't. Asexuality is not celibacy - celibacy is a choice to abstain from sexual intimacy while asexuality is an orientation which results in lack of sexual attraction. Welcome to Yahoo's oldest club for asexual people and their friends.[29]

Here we see asexuality defined in terms of sexual attraction. Diversity among asexuals is stated, specifically regarding romantic orientation and whether or not people "experience arousal." Further, asexuality is said to be an orientation and it is distanced from celibacy.

---

[28]Retrieved 2012-12-28 from http://web.archive.org/web/20050329234356/http://groups.yahoo.com/group/havenforthehumanamoeba/

[29]Retrieved 2012-12-28 from http://web.archive.org/web/20060618000138/http://groups.yahoo.com/group/havenforthehumanamoeba/

"The nonsexuality rant" was periodically updated over time. In late 2005 or early 2006, a section was added to the beginning, part of which reads "I'm not interested in sex–call it nonsexuality, asexuality, antisexuality, whatever, it amounts to this: I have no sex drive." In late 2007, the title was changed from "Nonsexuality rant" to "Asexuality Top 10." Sometime between April 2009 and August 2010, the above quoted sentence was modified: "I'm not interested in sex–call it nonsexuality, asexuality, antisexuality, whatever, it amounts to this: I have no interest in sex, no sexual attraction to others, and no sex drive. (Asexuals are defined by most as people who don't experience sexual attraction, for the record.) In these changes, we see movement towards "asexuality" rather than "nonsexuality." In the last change, there is a move away from defining things in terms of not having a sex drive (given the idea of some asexuals having an undirected sex-drive) and towards a sexual-attraction-based definition, the parenthetical comment indicating that her not having a sex-drive is a feature of her asexuality, not of asexuality in general (as defined by most).[30] The opening of the most recent version distinguishes asexuality from celibacy: "Asexuals are people who aren't attracted to other people in a sexual way. It isn't a synonym for celibacy or a moral statement or a decision people make; it's an orientation, like being straight is."[31]

The third line of archival evidence for the AVEN shift is that the two relatively early online asexual communities I know of which took positions contrary to AVEN ended up closing down. The first is OFA. The earliest archived version of the OFA's FAQ has a somewhat unclear statement about masturbation, [32] but sometime between October 2001 and October 2002, the FAQ was greatly expanded. The first question and answer were about masturbation: "...If I called all the shots, I would say no, someone who masturbates isn't antisexual. I see masturbating as a sexual activity, which pretty much goes against the definition of antisexualism. But I'm sure there are people out there who prefer themselves to other people..." [33] The second question/answer pair appears to deny the concept of

---

[30]Retrieved 2012-02-28 from http://web.archive.org/web/20100817230400/http://swankivy.com/writing/essays/philosophy/asexual.html

[31]Retrieved 2012-02-28 from http://swankivy.com/writing/essays/philosophy/asexual.html

[32]Retrieved 2014-03-19 from http://web.archive.org/web/20011019153340/http://www.geocities.com/WestHollywood/Village/1563/faq.html

[33]Retrieved 2014-03-19 from http://web.archive.org/web/

romantic orientation by treating "attraction" and "sexual attraction" as synonyms.[34] The third question is about whether there is a difference between being asexual and antisexual, and the answer suggests there is not. Given the date of these changes and its explicitly mentioning of the term asexual, this new FAQ is likely responding to AVEN and/or the LJ asexuality community, and (at least partly) rejecting some of their ideas.

An early online asexual space that was created somewhat after the AVEN forums at first called itself The Official Asexual Society.[35] Its founder had been involved in asexual visibility for some time, but had a rather different understanding of asexuality than AVEN, as the page on "What asexuality is NOT" makes clear:

> Some people think they are Asexual as they seem to poses a Sex Drive, but never associate those feelings with other people. They may get the mechanics of sexual urges, but they never desire to act on them to have sex. This may vent itself in many ways, but it seems masturbation is a common one.

> "Masturbaters, or Mono Sexual", You can have a sex drive without it actually going anywhere. One has the mechanics of sexuality and can even become sexually aroused. A person who masturbates is not an Asexual.[36]

Sometime between August 2004 and April 2005, the site changed its name to "The Official Nonlibidoism Society. While the website does not appear to give any reason explaining this change, according to the AVENwiki's page on nonlibidoism,

> Due to the popularity of a more inclusive definition of asexuality, the Official Nonlibidoism Society ceased to use the term 'asexual' for its members, believing that it had 'by now become almost synonymous for solo-sexual [or] masturbator' (reference no longer available). Nonlibidoism was a much more stringent definition

20021008114717/http://www.geocities.com/WestHollywood/Village/1563/faq.html

[34]"An antisexual or asexual by definition doesn't have a sexual attraction to anything, so no they aren't attracted to the opposite sex. However, those who are homosexual or bisexual are." Retrieved 2014-03-19 from
http://web.archive.org/web/
20021008114717/http://www.geocities.com/WestHollywood/Village/1563/faq.html

[35]Retrieved 2012-12-29 from http://web.archive.org/web/20030818023940/http://www.theofficialasexualsociety.com/productssimple.html

[36]Retrieved 2014-03-19 from http://web.archive.org/web/20040203001856/http://www.theofficialasexualsociety.com/productssimple6.html

than AVEN's standard description of asexuality, for people who have no sex drive, and have never had one.[37]

The Official Nonlibidoist Society seems to have had forums that closed down several years ago, but it is difficult to find reliable information. According to Titman (2012), they had a write-in survey people could fill out to find out if they were asexual or not. Then, the site's owner would decide if that person was asexual (by a rather stringent standard), and if so, joining was permitted. While I doubt that the Official Asexual Society was ever a serious rival with AVEN for dominance in asexual discourse, it serves as a foil of their own approaches in both Jay's (2006) and Titman's (2012) telling of asexual history.

A fourth line of qualitative evidence that the AVEN shift has occurred is a recollection from Jay's (2006) presentation about the history of asexuality, drawing on personal experience and memory, the WayBack Machine, and their own (well-kept) email archive:

> We had this small community, we had a label-human asexual-that we could find each other from. But even with that, there was still very little agreement about what being asexual meant. Rather than using the lens of sexual orientation to frame our experiences, [on] most asexual mailing lists and discussion groups of that era [...] it was basically the experience of not having sex. It was about people who didn't have sex. And they didn't have the language that we have today. There was no clear definition.[38]

Here it is explicitly claimed that there has been a shift, and three aspects of what things had been like previously are highlighted, aspects that Titman's audience, largely made up of asexuals, was expected to understand as contrasting with how asexuality was understood at the time of the presentation. First, in early asexual communities, there had not been a "clear definition" or "agreement about what being asexual meant." Second, the understanding of asexuality has shifted from being primarily about behavior (i.e. celibacy) to being a sexual orientation. Third is "the language we have today," indicating that a major part of the shift

---

[37]Retrieved 2012-02-27 from http://asexuality.org/wiki/index.php?title=Nonlibidoism
[38]The transcription is my own.

that has occurred is linguistic. Additional lines of evidence to investigate whether there has been an AVEN shift will be discussed in Chapter 6.

4.5.2 The rise of intermediate categories

My second hypothesis concerning the development of asexual language is that in recent years there has been a rise of intermediate categories. That is, there has been an increase in usage of–and convergence on–terms that refer to people somewhere between being asexual and sexual/nonasexual.[39] While I have not found any reference to terms for such individuals in the early static content sites, there has been discussion of this on AVEN for some time. Some "early" terms I have encountered include "hyposexual," "semisexual," and "pseudosexual," with hyposexual being the most common. However, two different terms have come to prominence in recent years. The first is demisexual, as well as the clipping demi. The second is variously gray-asexual, gray-A, graysexual, grace[40] etc.[41] Subsequent to these terms coming into more widespread usage, the morphemes *gray-* and *demi-* were applied to the concept of romantic orientation, resulting in terms such as gray-aromantic, gray-romantic, and demi-romantic.[42]

The term "demisexual" was originally intended with a rather technical meaning that requires some explanation. It was coined in 2006 by the AVENite (this is the standard term for members of AVEN) sonofzeal (John Vipperman).[43] Shortly after joining AVEN, he was introduced to a model of sexuality that the AVENwiki presently calls the "Primary vs. secondary sexual attraction model": Under this model, primary sexual attraction is an instant attraction to people based on instantly available information such as their appearance or smell which may or may not lead to arousal or sexual desire. Secondary sexual attraction is considered to be an attraction that develops over time based on a person's relationship and

---

[39]As of writing, the most commonly used out-group terms in asexual discourse are "sexual", "non-asexual", and "allosexual."

[40]This is a blend of "gray" and "ace," where "ace" is used as a clipped form of "asexual."

[41]For most of these, the spellings "gray" and "grey" are both attested. I will generally follow American orthographic norms, but original spellings will be preserved in quotes.

[42]For many of the neologisms discussed in this paper, the use of hyphens varies.

[43]Throughout this dissertation, all names and/or screen names are used with permission of the individual, except in cases where they have already made their name very public.

emotional connection with another person....Most sexuals in romantic relationships feel both primary and secondary sexual desire. The term demisexual, under this model, tends to refer to people who experience secondary sexual attraction but not primary sexual attraction.[44]

The terms gray-A and gray-asexual both first appear in the same thread from April 2006. A post from October 2006 by the AVENite Hallucigenia well illustrates the state of relevant asexual terms at the time. Hallucigenia was later a member of AVEN's Project Team for several years, and helped to popularize the gray-A terminology:

> All our hyposexual, demisexual, semisexual, pseudosexual, kinda sorta sexual, and just plain confused people should represent. (I'm going to use "hyposexual" as an umbrella term throughout, because it seems to be the most widely recognized one, but if anyone has a term that they prefer, then please post and use/explain it!)[45]

The responses in the thread indicate that many felt that this was an important topic, some said they had never heard of "pseudosexual," , and it was in response to this post that Hallucigenia was first introduced to the term gray-A by the person who first coined it. Further, this post indicates that although the concept existed, there was no generally accepted terminology for it, but the mostly commonly used term seemed to be hyposexual.[46] This contrasts with the present state of affairs as seen by the following evidence.

In September 2011, AVEN created a new forum called "The Gray Area" specifically for discussion of these sorts of topics. In Spring 2012, a static-content "Grey-A and Demisexual FAQ and link thread" was pinned and locked in that forum.[47] Below are the first three question/answer pairs for its gray-a FAQ:

> **What is a grey-a?** A grey-a is a part of the sexuality spectrum that is close to asexuality. A grey-a may experience sexual attraction under limited circumstances

---

[44]Retrieved 2014-04-18 from http://www.asexuality.org/wiki/index.php?title=Demisexual

[45]Retrieved 8/5/2012 from http://www.asexuality.org/en/index.php?/topic/19379-hyposexuals-semisexuals-demisexuals-pseudosexuals-unite/

[46]This impression is confirmed by quantitative data. See Chapter 5.4.1

[47]In many online forums, threads are listed by chronological order based on the most recent post in that thread, meaning that whenever someone posts in a thread, that thread is moved to the top of the forum. The exception to this is "pinned" threads, which remain at the top of the forum above all non-pinned threads. If there are multiple pinned threads, they are listed in order of the most recent post. If a thread is locked, it cannot be posted in (unless it is unlocked, which can only be done by someone moderator powers in that forum.)

or to a lesser degree than sexual people.

How is a grey-a different from an asexual? Many grey-as still experience sexual attraction. An asexual is defined as a person who does not experience sexual attraction.

How is a grey-a different from a sexual? The grey umbrella is very large and the answer to this question is very complicated. Each grey may have varying reasons ranging from not experiencing sexual attraction to the same degree as sexual people, not experiencing sexual attraction as often, or something completely different.[48]

Given their position in a pinned FAQ, these views are presented as being authoritative. Further, they would not have been pinned and remained so unless their content was acceptable–or at least reasonably unobjectionable–to the others managing the site and its static content.

Another piece of evidence that there has been a rise of intermediate categories concerns the term ace. My first experiences with this tem in asexual discourse were as a sort of slang term for an asexual. At some point in summer/fall 2011, I first observed it being used as a superordinate category for asexuals, demisexuals, and gray-a's, sometimes occurring in the expression the ace spectrum. Zwicky (2012) made a blog post about ace in asexual discourse, noting that it is a somewhat anomalous clipping: *heterosexual*, *homosexual*, and *bisexual* all have clipped forms (*hetero*, *homo*, and *bi*), but unlike ace for asexual, all of these clippings are prefixes. One commenter reported that at the San Francisco Pride Parade in 2011, David Jay said this is how he understood it, and according to another commenter, this meaning of ace had been used before then, but following David Jay's comments is when this meaning became much more common. A felt need to have a superordinate term for asexuals, gray-A's, and demisexuals suggests a perception that the latter two groups were playing an increasing role in asexual communities.

---

[48]Retrieved 2012-12-29 from http://www.asexuality.org/en/topic/74042-grey-a-and-demisexual-faq-and-link-thread/

4.5.3 Romantic Orientation

For a qualitative account of romantic orientation terminology, I will primarily rely on the AVENwiki. After the AVENwiki was started in September 2006, one of the earliest articles to be added was the Lexicon page[49]. In the very first version, there were 7 items listed, including *aromantic*, *biasexual*, *heteroasexual*, and *homoasexual*. This suggests that, at the time, these were especially salient as important terms in AVEN discourse that are not commonly used elsewhere.

By the end of the day, there were several more edits. These included a) replacing *bi-*, *hetero-*, and *homoasexual* with hyphenated forms, b) adding *bi-romantic*, *hetero-romantic*, *homo-romantic*, *gay-A*, and *straight-A*.

The explanation for *hetero-asexual* was "being romantically attracted to people of the opposite sex, but not sexually attracted; or someone for whom this is true." (*straight-A is said to be a synonym*) By contrast, *hetero-romantic* was described as ": someone who is romantically attracted to people of the opposite sex (and may be sexual or asexual)". The explanations were analogous for the *-asexual* and *-romantic* words with the *bi-* and *homo-* prefixes, and also for *gay-A*. However, the explanation for *aromantic* is not parallel, nor does it even mention attraction. Rather, it was described as meaning "not interested in romantic relationships". By contrast, a page for *aromantic* was created in November 2006, which states, "There is no concrete definition for 'aromantic'. Some people think of it as the lack of romantic attraction, whereas others define it as the lack of desire to be in an exclusive, romantic relationship.".

This suggests that, by mid-2006, there was no obvious preference for the *X-asexual* forms versus the *X-romantic* forms. A change in June 2008 suggests that, by then, the *X-romantic* forms had come to dominate: The explanations for the *hetero-asexual* was changed to "a shorten form of hetero-romantic asexual" (and likewise for *bi-* and homo-asexual).

Beginning in October 2010, the AVEN Project Team started an initiative (beginning by creating a thread on AVEN) to help get more people working on updating the wiki, and

---

[49]http://asexuality.org/wiki/index.php?title=Lexicon&oldid=1419

to make some decisions about what sorts of changes to make. In the opening post, the question was asked "Are there any pages that you think need to be updated a lot?" In a subsequent post,[50] the same PT member summarized responses: "Aromanticism. The lexicon in general". Based on views expressed in that thread, *gay-A*, *straight-A*, and *X-asexual* terms were moved to a newly created section called "Dated asexual terminology", and the explanation of, e.g., *hetero-asexual* was changed to " a shortened form of the more commonly used term, heteroromantic asexual."

4.5.4 Repulsed and Indifferent

In Chpater 1.2, I proposed using the terms *undesire* and *nondesire* as the contrary and contradictory of *desire*, respectively. In this sense, *nondesire* means "not desire to X" whereas *undesire* means "desire to not X". Both exclude "desire to X", but *nondesire* leaves open the possibility of indifference. As an identity, asexuality generally involves an assumption of not being interested in sex. As such, the difference between *undesire* and *nondesire* is the possibility of neutrality with respect to desiring sex. In asexual discourse, this distinction is sometimes expressed in terms of being *repulsed* vs. *indifferent*. As with romantic orientation, this section will rely primarily on the history of the AVENwiki.

A wiki page for *repulsed* was created in October 2008, and another for *indifferent* was created on the same day. Each is quite short, and neither has been revised much since then. I quote here the opening of each:

> Indifferent is a term used by some asexual individuals to indicate that they feel nei-ther revulsion toward nor powerful desire to engage in sex, as in, "I'm an indifferent asexual" or simply "I'm indifferent."[51]

> Repulsed is a term used by some asexual individuals to indicate that they find sex disgusting or revolting, as in, "I'm a repulsed asexual" or simply "I'm repulsed."[52]

Both terms were added to the main lexicon page when the page was undergoing major revisions in October 2010, based on a recommendation in the thread about updating the

---

[50]Retrieved 2016-05-29 from http://www.asexuality.org/en/topic/55678-updating-the-wiki/
[51]Retrieved 2016-06-01 from http://asexuality.org/wiki/index.php?title=Indifferent
[52]Retrieved 2016-06-01 from http://asexuality.org/wiki/index.php?title=Repulsed

AVENwiki. The number of page views is relatively low for the pages for *indifferent* and *repulsed*. As of May 2016, *indifferent* had been accessed 8,745 times,[53] and as of April, the page for *repulsed* 31,587 times.[54]. (By way of contrast, the page for *Gray-A* had been accessed over 600,000 times; *aromantic* over 450,000 times.) Although neither page is very popular, *repulsed* had been accessed more than three times as often, which suggests a difference in importance of the two terms.

The inclusion of *repulsed* and *indifferent* in the AVENwiki, suggests that these are important terms in asexual discourse, but the evidence appears much weaker than for many of the other terms discussed. So in May 2016, I performed several searches via Google for other webpages with lists and definitions for asexuality-related terms. Other than the AVEN wiki, I found four.[55] All four have some form of *repulsed*, but only one of them (Anagnori) includes a form of "indifferent" (*sex-indifferent*). The different lists differ on what form of *repulsed* they use. Three use *sex-repulsed*, and one uses *repulsed*.

Taken together, these suggest that *repulsed* and *indifferent* are both important concepts in asexual discourse, but *repulsed* is more salient than *indifferent*. They also suggest that further investigation should look both at forms where "sex" is part of the wider context, and forms where it is part of a hyphenated word. The quotes from the AVENwiki also suggest looking for the terms *repulsed asexual* and *indifferent asexual*.

4.6   Conclusion

In this chapter, I examined the growth of online asexual communities, the conceptual options used in their vision of asexuality, and the development of some important terms and concepts.

In Chapter 3, I found that the conceptual building blocks for something like current asexual discourse can be found in academic literature since at least the end of the 19*th* century.

---

[53]https://web.archive.org/web/20160528224140/http://www.asexuality.org/wiki/index.php?title=Indifferent

[54]https://web.archive.org/web/20160422184315/http://www.asexuality.org/wiki/index.php?title=Repulsed

[55]1) Asexuality Archive. "Glossary": http://www.asexualityarchive.com/glossary/
2) Asexuals Anonymous. "a glossary of asexual and aromantic-related terms": http://asexualsanonymous.tumblr.com/glossary
3) Anagnori's "Words and concepts used in asexual communities": http://anagnori.tumblr.com/post/67669933207/words-and-concepts-used-in-asexual-communities
4) Matthew's Place. "Master List of Asexual and Aromantic Vocabulary". http://www.matthewsplace.com/voice/asexual-and-aromantic-vocabulary/

In this chapter, I began by examining some popular-level discourse about asexuality in US and UK newspapers in the 1950-1990s. These revealed that the basic conceptual building blocks for asexual-identity had existed in popular discourse for at least several decades prior to the emergence of online asexual communities. Further, they provided evidence of at least some people who identified as asexual and thought of this in terms of sexual orientation. However, the available evidence is extremely sparse.

Following this, I gave an overview of the development of online asexual organizing, tracing it from mostly static content sites in the 1990s, to growing dynamic content in the early 2000s, to the dominance of AVEN beginning around 2003, and then the considerable growth in asexual discourse elsewhere on the internet, especially since around 2008.

I argued that qualitative historical evidence suggests two major shifts in online asexual discourse, which I am calling the AVEN shift and the rise of intermediate categories. In addition, I also briefly examined two other groups of asexuality-related terms: romantic orientation terms, and the terms *repulsed* and *indifferent*. In addition, I found qualitative evidence for a shift away from *X-asexual* to *X-romantic* terms. In looking at the terms *repulsed* (or *sex-repulsed*) and *indifferent* (or *sex-indifferent*), I found qualitative evidence that both *undesire* and *nondesire* are important parts of asexual discourse.

In the next chapter, I discuss the creation of asexual and non-asexual corpora to further investigate these matters through quantitative methods.

Chapter 5

Asexual Corpora

The first research question addressed by this chapter is whether one can even talk meaningfully about (written) "online asexual discourse". Is there anything that distinguishes this from other kinds of discourse? An affirmative answer to the first question is a logical prerequisite for investigations of how it might be different from non-asexual discourse (however asexual and non-asexual discourse are operationalized). In theory, a great many kinds of differences could be relevant for distinguishing asexual discourse from non-asexual discourse. Even restricting consideration to online (mostly) text-based discourse, this could include individual words, collocations, orthography, syntax, in-group humor, concepts, non-linguistic community norms, etc.

From my experience participating in a number of online asexual communities, and lurking on several more, the most salient features of online asexual discourse involve terminology, in-group jokes, common experiences among asexuals, and other sorts of assumed background knowledge. Community norms for online asexual communities vary from community to community, and these are largely inherited from norms for the platform of the community (i.e. much of AVEN's community norms are inherited from norms for forums, asexual blogging community norms generally derive from blogging norms more generally, etc.) For these reasons, looking at lexical differences is a good place to test whether asexual discourse can reliably be distinguished from non-asexual discourse. Further, this is consistent with the largely lexical focus in my main research questions.

After providing evidence for an affirmative answer, I move to my other main research questions: What words are most characteristic of asexual discourse, and what can these tell

us about asexual discourse? Are expressions of undesire and/or nondesire important parts of online asexual discourse? Does quantitative research support the *AVEN shift* and *rise of intermediate categories* hypotheses?

In order to answer my research questions, this chapter and the next primarily use two main kinds of data: corpora based on online asexual discourse and an online survey. A major limitation of using corpus data is that, like many online communities, the asexual communities the corpora are based on, are "demographically lean"(Iorio, 2011).[1] While there is a great deal of data, it is difficult to study linguistic variation with the standard sorts of variables used in variationist work, such as age, gender, or socio-economic status. By contrast, the online survey was able to get better demographic data about survey-takers, but survey data provides information on self-reported usage, which is not always consistent with actual usage. Further, the only kind of diachronic data that is possible through the survey is retrospective self-report. These limitations of the survey data are among the major strengths of using corpus data. These two approaches well-complement each other.

The rest of this chapter is organized as follows. In Section 5.1, I will discuss the creation of four corpora: The AVEN-core corpus (containing most of AVEN's forums viewable to non-members), the AVEN-other corpus (containing AVEN forums that are not generally for discussing asexuality-related topics), the Asexual-other corpus (containing asexuality blogs and some asexual communities other than AVEN), and a reference corpus, which was created from forums on the xkcd website.[2]

In Section 5.2, I will discuss the use of a Naïve Bayes classifier to test whether asexual discourse can be distinguished from non-asexual discourse. The training data for asexual discourse come from the AVEN-core corpus and the non-asexual training data come from the reference corpus. A source of information that, intuitively, seems likely to artificially inflate accuracy is names for users. Methods for removing these are discussed briefly below, and

---

[1]The online communities for which I have data are not as demographically lean as on many MMORPGs (massively multi-player online role-playing games), such as the one that Iorio studied. However, extracting relevant information (e.g. standard sociological variables such as age, gender, level of education, etc.) is likely impossible for many users in my data, and for many extracting it would be a non-trivial computational problem beyond the scope of this dissertation.

[2]forums.xkcd.com

in more detail in Appendix A. Results indicate moderate distinguishability at the level of individual posts when only considering individual words, even after excluding names for users and AVEN-specific terminology. Because it is possible that this identified AVEN discourse vs. xkcd discourse, I compare ran the classifier on the AVEN-other and Asexual-other corpora. Results for the Asexual-other discourse are similar to results for the AVEN-core, while results for the AVEN-other corpus do not (although they do not pattern with the xkcd data either). This suggests that the classifier is identifying xkcd discourse vs. asexual discourse (rather than AVEN discourse).

Then in Section 5.3 I do a keyword analysis to see what words and themes are especially characteristic of asexual discourse.

Of the keywords, I select three groups of terms for analysis of usage over time: intermediate category terms (Section 5.4.1), romantic orientation terms (Section 5.4.2), and *repulsed* and *indifferent* (Section 5.4.3). Results for intermediate value terms provide strong confirmation of the *rise of intermediate categories* hypothesis. Results for the other two also point in the same direction as the qualitative analyses in Chapter 4.5.

5.1 Corpus design

In this section, I will discuss the construction of the AVEN-core, AVEN-other, and Asexual-other corpora and of the reference corpus.

5.1.1 The AVEN corpus

The AVEN corpora consist of data from most of the asexuality.org/en that are viewable to non-members[3], from when the AVEN forums were created in 2002 through late November 2013 when I obtained the final version of the data.

Most of AVEN's forums are for topics relating to asexuality in some way, but there are several that are not generally about asexuality. For example, the forum "Just for Fun" (JFF)

---

[3]Excluded are "Alternate Language Forum" because of its high rate of non-English language content, and "Community Link Forum" because it has no threads.

is described as follows: "Also known as the Mudkip Empire. Where the hard-postin' folks of AVEN come at the end of a long day to get low-down and goofy."[4] Likewise, the description of Off-A is "Talk about entertainment, hobbies, every day life and all non-asexuality related things here!" Because of this, a distinction between "core" forum (e.g. ones that constitute "asexual discourse") and more peripheral ones is warranted. This requires a binary-decision task for all AVEN forums in my data.[5] In my view, most forums clearly belong to either the "core" or "peripheral" group, although there are some unclear cases. The primary criteria for determining whether a forum was "core" or "periphery" are as follows:

- It must have at least 500,000 words.[6]

- It must be largely about (a)sexuality or a closely related topic (i.e. romance, gender): This is established from the forum name, its description, and (if necessary) examination of thread titles.

Of the 22 forums in my AVEN data, 16 are included in the AVEN-core corpus. These are listed in Table 5.1, along with AVEN's description of the forum and its word count. The remaining 6 forums are grouped together in the AVEN-other corpus. These are listed in Table 5.2, along with AVEN's description of each forum and its word count.[7] The AVEN-core corpus has approximately 74 million words, and the AVEN-other corpus approximately 37 million.

From the names, descriptions, and size, inclusion in the AVEN-core corpus or in the AVEN-other corpus is straightforward, but some explanation is warranted in a few cases. To begin with ones included in the AVEN-core corpus, "Visibility and Education Projects" is closely related to the *VE* in *AVEN*, and it is often used for discussion of asexuality-related projects people are doing, are thinking about doing, or would like to see someone

---

[4]Retrieved 10/9/2014 from http://www.asexuality.org/en/

[5]The "Alternate Languages" forum is excluded entirely because of the large proportion of non-English language content.

[6]Two of the 22 AVEN forums–Site Info Center and World Pride 2012–each have less than 100,000 words, while all other forums have over 1,000,000 words. Their z-scores for the log transformed number of words are z = -3.07 and z = -2.01, while all other 20 z-scores fall in the interval (-0.43, 1.35)

[7]The method that I used for scraping the data first obtained the url's for each forum to be scraped. Then it went page-by-page through the index for each forum, making a list of url's, and these were each saved in a different file. Then each thread was scraped. Because the urls were obtained by going from the first to last page in a forum's index, it is possible that a small number of threads were missed due to changes made while obtaining the thread url's.

Forum	Description	Word countte
Announcements	A place for forum-wide announcements, calls for interviews and the like. Make sure to read up to see what's new.	1,037,177
Welcome Lounge	Drop in and say hi, tell your story and get to know other newbies and AVEN members	9,167,575
Asexual Musings and Rantings	Come here to talk, discuss, gripe, or revel about asexuality. Don't be shy!	20,879,868
Asexual Q&A	Questions about asexuality? Here's the place!	11,472,076
Asexual Relationships	A place for all discussions about relationships, be they romantic, friendly, familial, or anything else. Can you handle it?	9,800,678
Older Asexuals	A forum mainly geared toward discussion of issues facing the more mature asexuals, but open to everybody.	2,905,973
For Sexual Partners, Friends and Allies	Are you a friend, relative or romantic partner of an asexual? Have some questions or need some support? Then this is the forum for you! You're welcome to use the rest of the board, as well, of course.	3,553,007
Visibility and Education Projects	A place to bring all your visibility and education projects to the table.	1,140,595
World Watch	A place to post all media, research, and literature regarding asexuality, including our own!	1,221,329
Census Forum	Polls that give the statistics that make up AVEN. Silly polls should still go in JFF. If you think a poll somewhere else belongs here, please notify an admin or mod.	1,765,607
The Gray Area, Sex and Related Discussions	Asexuality and sexuality are not black and white concepts; come here to discover and discuss the grey area in between, demisexuality, and all things related to sexuality and sexual identities!	1,203,957
Romantic and Aromantic Orientations	A place to discuss romanticism, aromanticism, the area in between, and the many kinds of attractions.	1,125,084
Gender Discussion	Gender is the agenda here. Feel free to discuss anything relating to gender in all of its forms.	4,155,032
Celebration Time!	A forum for sharing your accomplishments, kudos, and whatever else might be making you feel great about your life. If you're happy and you know it, post in here!	1,145,236
Site Comments	Post any questions, comments or suggestions you have for the AVEN website.	1,928,344
Admod Archives	For all your modly needs! Declassified threads are added regularly, so check in often if you are curious about what your Admods are doing behind the scenes.	1,149,279
Total		73,650,817

Table 5.1: Name, description, and word count for the forums in the AVEN-core corpus

Forum	Description	Word count
Off A	Talk about entertainment, hobbies, every day life and all non-asexuality related things here!	14,375,316
Just for Fun	Also known as the Mudkip Empire. Where the hard-postin' folks of AVEN come at the end of a long day to get low-down and goofy.	13,003,956
Philosophy Politics and Science	A place to discuss current events, ask questions, and share knowledge.	7,901,118
Open Mic	A place for members to discuss all aspects of the creative processes, in both theory and application, and to showcase and improve their awesome talents	1,776,802
Worldpride 2012	A forum for the planning of Worldpride 2012, a large international visibility effort, which will take place in London on 7 July 2012.	70,059
Site Info Center	The place to find the rules, how-to, and other wordy but useful documents	11,564
Total		37,138,815

Table 5.2: Name, description, and word count for the forums in the AVEN-other corpus

do. Projects ready for announcing to a larger audience go in Announcements, as do (most) calls for participants for research (and sometimes feedback) for research on (or including) asexuality, and calls for interviewees for journalists (writing articles about asexuality). "Site Comments" and "Admod Archives" could probably go either way, but my preference was to put these in the AVEN-core because these often include discussions about asexuality, asexual communities, and what sort of site AVEN should be to accomplish its goals.

In two cases, assignment was especially unclear: "Celebration Time" and "Open Mic".[8] To use an empirical basis for assignment in these two edge cases, the decision was based on the accuracy of the bag-of-words classifier when neither was included in the training data. If the accuracy was at least 70%, it was included in AVEN-core (Celebration Time!), otherwise it was not (Open Mic).[9]

---

[8]Examining the thread titles for Celebration Time, most deal with members' birthdays, their "AVENiversaries" (anniversary of joining AVEN), or achieving a noteworthy milestone in the post-count. They also often involve positive experiences relating to asexuality, such as finding a significant other, getting married, having a positive experience coming out to someone as asexual, going to an asexual meet-up, etc. And many are about other sorts of positive experiences, like getting a job. Open Mic is for posting creative works such as poems or stories. Sometimes these related to asexuality, but usually they do not.

[9]Using accuracy to establish what to include in the AVEN-core may seem circular—and it would be if this were the primary criterion for classification. However, this criterion was used only for two forums (comprising about 2.6% of the 110.8 million words in the combined corpora) to provide an empirically-grounded basis for classifying forums where the primary criteria were inconclusive.

5.1.2 Other asexual corpora

While AVEN was, for a number of years, the largest online asexual community by far, it is not the only locale for online asexual discourse. Therefore, additional asexual corpora were constructed to supplement the AVEN data. Various factors were involved in deciding what to include:

- Preference was given to earlier communities, both to help investigate whether there has been an AVEN shift and to add data to the time period when the AVEN data is most sparse.

- The language had to be primarily in textual form (i.e. podcasts and YouTube channels were excluded).

- Because the scripting required varied from locale to locale, preference was generally given to larger locales over smaller ones.

- Smaller locales were more likely to be included if the format closely resembled that of a larger one.

- The content had to be viewable without a login at the time it was posted.

To determine what communities were considered to be relevant, I examined AVEN's links over time via the WayBack machine, as well as discussion on AVEN about what other asexual communities existed. I included relevant LiveJournal communities if they had at least 50 entries (n=2), and I included forums if they had at least 500 members (n=1 other than AVEN). Using these criteria, I have compiled corpora from the Yahoo! Group Haven for the Human Amoeba (only using posts from 2000 through June 2005), the LiveJournal communities Asexuality and Asexuals, and the forum Apositive. For Apositive, all forums viewable to non-members were included with the exception of "General Discussions and Fun."

In addition to these fairly well defined asexual communities, I constructed an Asexual Blog Corpus, and I restricted this to blogs not on Tumblr. One reason for constructing an asexual blogging corpus is that, prior to the large role of Tumblr, an "asexual blogging community" was one of the main online asexual "communities" other than AVEN. (I put "communities" in quotes because, in some cases, there was clear interaction between different bloggers, while

Category	Words
Blogs	2,770,839
LJ Asexuality	2,574,798
LJ Asexuals	130,924
Haven for the Human Amoeba	438,253
Apositive	703,062
Total	6,617,876

Table 5.3: Word count for the asexual-other corpus and its subparts

other blogs were less connected.) Tumblr was excluded largely for pragmatic reasons: Its format is quite different from other blogs, it would be difficult to create a defensible criterion for inclusion/exclusion in a Tumblr asexuality corpus, and the complexity of the technical aspects require that studying asexual discourse on Tumblr be left for future research.[10]

Of necessity, criteria for inclusion in the blog corpus were somewhat more subjective than for the previous asexual corpora. Blogs for potential inclusion were identified in part through prior knowledge of asexual blogs, the list of blogs listed on the AVENwiki's page about asexual sites, and the blogrolls of blogs identified through these means. Blogs were included in the corpus if they were primarily about asexuality and closely related matters. Some blogs did not fit this criterion, but did have a specific subsection about asexuality marked with tags. In these cases, the relevant subsection was included. In addition, I sent my list to two people who are presently active in the asexual blogosphere to see if I had missed anything that should be included. Blogs were excluded if clicking on a content warning was necessary to view content.[11] One blog was excluded because posts did not have a datestamp. Based on these criteria, 70 blogs were identified for inclusion or partial inclusion. The size of the non-AVEN asexual corpora is given in Table 5.3

For most of the blogs, both posts and comments are included in the corpus. All posts and comments are tagged with the date and a unique postID. Unlike the other corpora, no author field was included because of the many-to-many relationship between individuals who write the content and the name it is posted under.[12] In some cases, the blog had been

---

[10]In general, most blogs in the blog corpus use Wordpress, Blogspot, or LiveJournal. The considerable variation in format for Wordpress blogs proved difficult to deal with, and the difficulty of creating a Tumblr corpus would be well beyond this.

[11]Based on this, one blog was excluded. A second blog had two versions, one with a content warning and without one. Most comments were one the version with the content warning, but the copy without comments was included in the corpus.

[12]Technically, this is probably true of the other corpora, but to a much lesser extent. For the blogs, LJ communities, and

deleted by the author, but the most or all of the posts remained archived in the Google Reader archives, which were publicly available until mid-July 2013.

Information for a blog could be downloaded from the Google Reader archive, via a page by Agarwal (2013), as long as one had the url of the RSS feed (or former url, for deleted blogs). In one case, the earliest posts were missing from this and the Way Back Machine was used to retrieve these. Historically, including these is important because three of the most popular (and prolific) blogs from the years 2008-2010, which is when asexual blogging first became a sizable part of asexual discourse, were subsequently deleted by their authors. For blogs retrieved this way, posts–but not comments–could be included in the corpus.

5.1.3   Reference Corpus

To implement the Naïve Bayes classifier (Section 5.2) and to find keywords for asexual discourse (Section 5.3), it is necessary to have a reference corpus of "non-asexual discourse," which should be comparable to the main asexual courpus (AVEN-core). This relates to a basic difficulty that behavioral scientists must face: On the one hand naturalistic data gives us a good idea of how people actually behave, but because of the many uncontrolled variables involved, inferences about the role of any one variable are difficult to draw; while on the other hand, causal inferences can be more reliably drawn from experiments conducted in the lab where numerous variables are carefully controlled for, but this very fact makes the context somewhat "unnatural", thereby drawing into question the extent to which findings can be generalized to everyday life.

There is no website exactly like AVEN in every way, except that it is not about asexuality. Therefore, rather than trying to find "the perfect control", I aimed to find as good a control as reasonably possible. A major difficulty for this is that it is difficult to find internet forums with extensive "general discussion" because this kind of CMC (computer mediated communication) platform is primarily used for help concerning specific software applications, PC games, hobbies, religious groups, or sometimes professions. It is relatively uncommon to

the forums, authors may change their names over time, but on the forums, this is applied retroactively to earlier posts, whereas it is not applied retroactively on blogs. Also, guest posts are much more common on blogs than in the other places.

have a sizable subforum devoted to off-topic discussions.

After examining numerous forums for potential inclusion in the reference corpus, I have decided to use three forums from the forums associated with the webcomic xkcd (forums.xkcd.com): News & Articles, Serious Business, and General, as these were among the largest "general discussion" forums I could find. I believe this is a fairly good control for several reasons. First, it is forum data, so it should share features of asynchronous computer mediated discourse, and it should have basically the same format of posting and replying. Second, it covers roughly the same time period as the AVEN data: the xkcd data covers March 2007 – January 2014. Around 80% of the words in the relevant AVEN data are from this time period. A poll on xkcd about its users' age suggests a mean of around 17,[13] while data from the AAW11 survey suggests a median age of AVEN users of about 23 (p.c. Tristan Miller). The accuracy of these figures is uncertain, but they suggest roughly similar age demographics, with participants in online asexual communities being somewhat older.

Combined, these three xkcd forums have around 69.57 million words, which is relatively close in size to the 73.65 millions words in the AVEN core. Figure 5.1 shows the distribution of words per post for the two sets of data on a log-log graph–posts are ranked from longest to shortest. The x-axis represents a post's rank (i.e. for the 10th longest post, $x = 10$) and the y-axis is the number of words in that post. As can be seen, the distribution of words per post are very similar, with the exception of, approximately, the 0.3% longest posts–the xkcd forums appear to have a higher character limit, and there are more posts that approach this character limit.

5.1.4 Technical details

All scripts for constructing and analyzing these corpora were written in Python 2.7 (Python Software Foundation, 2010). In addition to the default libraries, I used BeautifulSoup (Richardson, 2013) and Mechanize (Lee, 2010) for retrieving and parsing webpages. Statistical analyses are performed using SciPy and NumPy (Oliphant, 2007), and graphs were

---

[13]Retrieved 10/22/2013 from http://fora.xkcd.com/viewtopic.php?t=44443&p=1757157
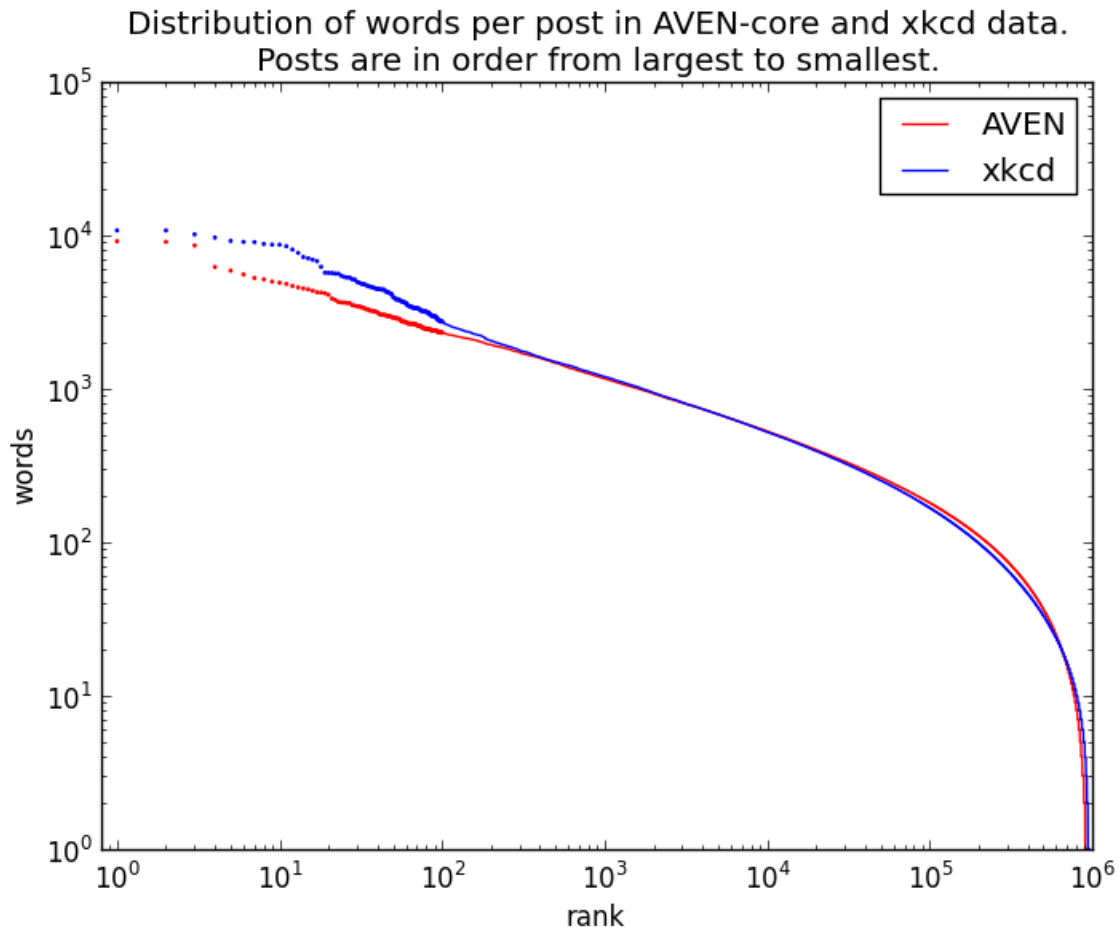
Figure 5.1: Distribution of words per post in AVEN-core corpus and in the xkcd data (reference corpus). For $x <= 20$ scatter-plots are used, and line graphs are used ffor $x > 20$.

created with matplotlib (Hunter, 2007). For all hypothesis tests, the smallest relevant α level I will use is 10^{-8}, and so p-values less than this will be reported as $p < 10^{-8}$. The reason for this is that p-values are a measure of the strength of evidence for rejecting the null hypothesis, and not a measure of the difference between groups.

In all corpora, Arabic numerals and the following characters

!@#$?.%^&*+"(),\[]_|:;<>~'="

were replaced with spaces, and the most commonly occuring other special characters were removed, with two exceptions: Word-internal hyphens and word-internal apostrophes were retained (non-ASCII apostrophes were replaced with their ASCII equivalent).[14] Multiple

---

[14]Some English words are standardly spelled with non-ASCII characters, so I did remove all such characters. Python 2.X is well-known for difficulties with working with unicode. Further, some of the blogs had multiple encodings in the same file.

Corpus	Words
AVEN-core	73,650,817
AVEN-other	37,138,815
Asexual-other	6,617,876
Reference (xkcd)	69,572,446
Total	186,979,954

Table 5.4: Word count for all four corpora

consecutive hyphens were assumed to be an em-dash, and so these were replaced with a space. Thus, *hetero-asexual* would be kept as *hetero-asexual*, while *hetero–asexual* would be treated as two separate words: *hetero* and *asexual*.

After all word initial or word final apostrophes and/or hyphens were removed (and if this did not leave the empty string), the first and last character of the string were checked against the unicodedata library's function *category* and the word was removed if either was classified as Lo (Other_Letter)[15] as these are primarily "letters" in non-Latin scripts.

For the corpora based on forums and blogs, material in blockquotes were removed, so that quoted material would not be counted multiple times.[16] In the corpora, boundaries between sentences (within a post) are ignored, while boundaries between posts are retained. Words containing "http" were removed prior to removing special characters. In corpora based on forums or LiveJournal communities, each post/reply is tagged with the author, the date, and a unique postID which is based on the thread's ID and the post's number within the thread. I also retained an uncleaned version of the corpora.[17] A summary of the four corpora is given in table 5.4.

There remains room for improvement with data-cleaning, but one version had to be declared final to proceed with analysis. Words with at least one character other than letters from a-z, apostrophes, or hyphens account for 0.011% of tokens in the core AVEN data and 0.024% of tokens in all of the asexual data, suggesting that improved cleaning would be unlikely to have a large impact on any of the findings based on the corpus data.

[15]For a list, see http://www.fileformat.info/info/unicode/category/Lo/list.htm

[16]For Haven for the Human Amoeba, the means of identifying quoted text was inconsistent, but generally involved < and >. Text identified as quoted text with these was removed.

[17]For communities other than AVEN, I have a copy of the html, whereas for AVEN, I only retained the uncleaned text, which already had material in blockquotes removed.

5.2 Naïve Bayes Classifier

Having constructed multiple asexual corpora, these can be used for a variety of purposes. I will begin by using a Naïve Bayes classifier to investigate three questions. First, can asexual discourse be fairly reliably distinguished automatically from non-asexual discourse? Second, if so, does the classifier generally classify even the peripheral parts of AVEN as "asexual discourse"? Third, does the classifier work for asexual communities other than the one used in the training data?

5.2.1 Basic description of Naïve Bayes classifiers

Naïve Bayes (NB) classifiers are a widely used method of machine learning that can be applied to a variety of problems, such as document classification, author identification, word sense disambiguation, among others. In my research, I will be using it as applied to document classification, and although it can be used to classify documents into several classes, I will only be using two classes, which I will label "asexual" and "nonasexual." In the discussion below, I will briefly give a general idea of how these classifiers work, with only enough formalism to explain the metric for Bayes factors in Section 5.2.2 below.

Very roughly, we want to give the classifier a document (which could be a single post, or a thread, or something else, depending on the goals of the user). The classifier will then use information from the text to establish which of the classes (i.e. asexual or non-asexual) it is most likely to have come from, and it then assigns the document to that class.

The key idea for a Naïve Bayes classifier is that, in order to make calculations easier and the running time faster, we make a major simplifying assumption about natural human language: The probability of each word occurring is independent of any other word in the text. That is, we assume that in language, there are no multi-word expressions, no grammatical dependencies, word order is totally irrelevant, and discourse context plays no role. For all calculations involved "The dog bit the man", "The man bit the dog", and "Man dog the the bit" are treated the same. Hence, the "Naïve" in "Naïve Bayes." This kind of

classifier is sometimes called a "bag of words" model because we can think of each class as being a bag of words, and documents are constructed from these by randomly pulling out a word from the bag, writing it down, putting it back in the bag, shaking the bag up, and then repeating the process a certain number of times. The classes differ in terms of the distribution of words in them (i.e. words that are used much more often in asexual discourse than non-asexual discourse have a much higher probability of being drawn from the asexual bag than the non-asexual bag).

Although it relies on a very simplistic model for language, Naïve Bayes classifiers have proved to be quite effective for a number of applications, and so they are used because they are relatively easy to implement, they have linear run-time, and they often work fairly well. Because of this, they often function as a baseline with which to compare other classifiers.

I will assume that, prior to looking at a document, it is equally likely to be asexual discourse as it is to be nonasexual discourse.[18] The classifier is used on "documents", where D is the set of possible documents, and $d \in D$ can be formally represented as a vector of attributes:

$d = (w_1, w_2, ..., w_n)$. For my classifier n is the number of words in a document, w_1 is the first word in a document, w_2 the second word, etc.[19] $C = \{asexual, nonasexual\}$ is the set of document classes. The bag-of-words classifier can be represented by the formula given in Equation 5.1,[20] where $NB : D \rightarrow C$, and $P(w_i|C = c)$ is the probability that, if a word is drawn from the "bag" for class c, it will be w_i.

$$NB(d \in D) = \arg\max_{c \in C} \prod_{i=1}^{n} P(w_i|C = c) \tag{5.1}$$

To calculate estimates for $P(w_i|C = c)$, for each class a set of training data is needed. The estimate for $P(w_i|C = c)$ is based on the relative frequency of w_i in c, with some

[18] That is, I am using a uniform distribution for the prior probabilities. According to Peng and Schuurmans (2003), the most common prior probability distributions used for naïve Bayes classifiers are the uniform and Dirichlet distributions. In reality there is vastly more non-asexual discourse than asexual discourse and so this is a somewhat naïve way of assigning prior probabilities in the present case, but it seems reasonable given the research questions the classifier is being used for.

[19] This approach is called a multinomial model (Lewis, 1998).

[20] For a derivation of this formula, see Peng et al. (2003).

modifications to deal with the problem of zero-probabilities (smoothing). The simplest method (additive smoothing) is simply to make a list of all words occurring in the training data for at least one class, then add some constant to the frequency for each word in each class, and then use the (smoothed) relative frequencies. When the constant is one, it is called plus-one smoothing. Although there are more complex approaches to smoothing that out-perform plus-one smoothing, I will use plus-one smoothing to keep the model as simple as possible.

5.2.2 Bayes Factors

Before discussing the implementation of the model, there is one metric used in my analysis that needs to be explained. Bayes Theorem, which is the basis for Equation 5.1, is given in Equation 5.2.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{5.2}$$

In Bayesian statistics, one interpretation of Bayes Theorem Equation 5.3 treats A as a hypothesis, and B as evidence. Using H (for "hypothesis") and E (for "evidence"), Equation 5.2 can be rewritten as in 5.3.

$$P(H|E) = \frac{P(E|H)}{P(E)}P(H) \tag{5.3}$$

The left-hand side, $P(H|E)$ is the posterior probability (i.e. the probability of the hypothesis in light of the evidence), and the right hand side is divided into to parts: $P(H)$ is the prior probability (i.e. the probability of the hypothesis before considering the evidence) and $\frac{P(E|H)}{P(E)}$ relates to the likelihood. A standard Bayesian interpretation of these three parts is that they provide a way of rationally updating (subjective beliefs about) whether some belief is true, in light of new evidence.

When we have two different hypotheses (or models), and we are wanting to test between these based on some observed data, the Bayes factor is a measure of how strongly the

evidence supports (or opposes) one of the hypotheses. (This means that the value of the Bayes factor depends not only on the evidence, but also on the hypotheses used.) For my research, the evidence/observation is either an individual word or all words in a document; the models/hypotheses are the two classes of text. For a word w_i, its Bayes factor (B_i) is as follows:[21]

$$B_i = \frac{P(w_i|c=Asexual)}{P(w_i|c=Non-asexual)}$$

If the Bayes factor equals 1, then the word provides no evidence for either class. If it is greater than 1, it provides evidence for classifying it as asexual. If it is less than 1, it provides evidence against this. In cases where the parameter under consideration is a continuous variable (e.g. an underlying variable that reflects the probability of success for a Bernoulli trial), calculating a Bayes factor can be relatively complicated. However, in the simplest case–e.g. when the parameter has only two values (as in the present case), then the Bayes factor is a likelihood ratio (Goodman, 1999; Jackman, 2009). Also, rather than using the Bayes factor itself, the measure of evidence that I will be using the log_{10} of the Bayes factor. The main advantage of this is that it creates a symmetry about 0 for evidence for vs. against a model. (i.e. a value of 2 predicts asexual discourse to the same degree that a value of -2 predicts non-asexual discourse). It also makes tables and graphs easier to read. The use of log_{10} rather than ln is also for readability (i.e. a log Bayes factor of 1 means that, according to the model underlying the classifier, the word/text is $10^1 = 10$ times more likely to have come from asexual discourse; a log Bayes factor of 2 means that it is $10^2 = 100$ times more likely to have come from asexual discourse, etc.)

There does not seem to be a widely used term for this measure (other than "log Bayes factor" or "the log of the Bayes factor"), although it has been used as a measure of the weight of evidence. e.g. Good (1950), who credits the idea to conversations with Alan Turing.[22] I will simply call it the log Bayes factor.

---

[21]Note the relationship with Equation 5.3: If the equation is replaced with two equations–in the first, H is replaced with H_1 (asexual) and in the second, H is replaced with H_2 (non-asexual)–and w_i is used as the value of E, then the Bayes factor is the ratio of the middle part, i.e. the part that is not the prior or posterior probability. Because $P(E)$ is the same for both H_1 and H_2, these cancel out when taking the ratio.

[22]The version that Good describes is slightly different that what I am using. Modeling it after the decibel scale, their metric was $10log_{10}(B)$. Also the term "Bayes factor" was not yet in use.

5.2.3 Details of applying a Naïve Bayes classifier

I will be using a multinomial Naïve Bayes classifier and use plus-one smoothing. Plus-one smoothing can cause serious problems in certain cases, such as when the training data is fairly small or when there is a very large relative difference in size between the training data for the different classes, and so I have tried to avoid these. In the (rare) case that a document was assigned equal probability for both classes, it was randomly assigned to one (with equal probability of being assigned to either).[23]

For the class "asexual discourse", I will use the "core" forums on AVEN. This choice is made because the AVEN corpus is far larger than the non-AVEN asexual corpora, and restricting the training data to AVEN data will enable me to see how well findings generalize to other asexual communities.

Another issue necessary for implementing the NB classifier is deciding on what is a "document". The most plausible options are to 1) use each post as a document, 2) use each thread as a document, or 3) some modification of one of the first two options. I wanted to have a large number of "documents" so I chose to use posts. Because poor performance on documents containing only a very small number of words would tell us very little regarding my research questions, I decided to use individual posts as documents, but place a lower limit on the word count–posts had to be at least 30 words to be used. Not wanting to throw out usable data on account of this, posts with less than 30 words were appended onto the following post. If these two combined were still less than 30 words, the combined post was appended onto the following post, etc.

Furthermore, to test my hypotheses, it is important to remove certain kinds of words likely to artificially inflate differences. Even if there were no meaningful differences (in terms of content or terminology) between the asexual data and the control data, it is conceivable that the classifier could perform reasonably well through the use of what is, intuitively, unimportant information. Perhaps the most obvious class of unimportant information is

---

[23]This situation only ever occurred with two documents, both from Just For Fun. One was 41 repetitions *imafail* and the other combines several posts from "the keyboard venting thread."

names for users. (This can include usernames, nicknames derived from these, or occasionally the person's actual name.)

In order to remove names for users, I developed two algorithms to automatically identify these. (Details may be found in Appendix A.) I ran these on data for AVEN and xkcd. For AVEN, I included both the core and peripheral data as names for users should be the same in both. Likewise, for the xkcd data, I used data from 22 public xkcd forums (although the 3 in the control data are by far the largest.)

After doing this, I made a list of the words that most strongly predicted the class "asexual" over "non-asexual" (i.e. I ranked words according to their log Bayes Factor, calculated based on all posts in the AVEN-core and the reference corpus) and examined by hand the top 500 items, and made a list of names for users in those items. If I recognized a word as name for a user, I included it in the list. If a word seemed likely to refer to a user, I used Google to search for the word on asexuality.org/en. I either added it to the list or not, depending on whether or not it referred to a user in the top result.

In addition to names for users, there are some other AVEN-specific words that might help to inflate the classifier's accuracy, despite not being an important part of asexual discourse more generally. For example, initialisms or blends referring to one of the AVEN forums (e.g. *jff*, *spfa*, *musirants*)[24] or some other part of the site (e.g. *avenwiki*) or groups of people on the site (e.g. *avenites*, *admods*, *pt*). They were manually added to the list of words to be removed if these were encountered in the list of the strongest predictors of AVEN. (AVEN-specific terms were coded separately from names for users, however.)

All strings tagged as a name for a user or an AVEN-specific term in any of these three methods was excluded. Further, if a string X was identified as a word for exclusion, then $X's$ was also excluded. Excluded words did not contribute to the minimum word count for documents.

---

[24] *JFF* is *Just for Fun. SPFA* is "Sexual Partners," Friends, and Allies. *Musirants* is "Asexual Musings and Rantings".

	Asexual Document	**Non-asexual Document**
Classified as asexual	89.09% (0.08)	10.91% (0.08)
Classified as non-asexual	17.84% (0.1)	82.16% (0.1)

Table 5.5: Mean (SD) accuracy as a % for 160 trials of the NB classifier with user names and AVEN-specific terms removed.

5.2.4 Results for the Naïve Bayes classifier for the testing data

To test the classifier, 160 trials were run on the data. In each trial, each document was randomly assigned to either the training data (75%) or the testing data (25%) using Python's "random" library. Because we know the actual source of each document used in the testing data, we can find what proportion of "asexual" the classifier assigned to the "asexual" group (true positives) vs. to the "non-asexual" groups (false-negatives).[25] Likewise, we can see what percentage of the "non-asexual" data the classifier classified as "non-asexual" (true negatives) vs. "asexual" (false positives).[26]

The means and standard deviations of the relative frequencies are given in Table 5.5. Note that because the right hand column is equal to one minus the left-hand column, the standard deviations for all cells in one row are the same.

Another way to examine the results of the classifier is to plot the distribution of log Bayes factors for the documents that it classified. In ten trials of the classifier, I calculated the log Bayes factor for each document in the testing data. The distribution of these is shown in Figure 5.2. (Note that documents to the right of the y-axis are classified as "asexual" and ones to the left are classified as "non-asexual".)

This evidence suggests that there are meaningful differences between asexual discourse and non-asexual discourse: Even when only lexical (relative) frequencies are considered and when "documents" are single posts (or a few short posts), there are enough differences between the asexual and non-asexual discourse data that the classifier can distinguish between

---

[25]There are a number of names for this metric including texitsensitivity, *hit rate*, *recall*, and *true positive rate*. For a summary, see http://en.wikipedia.org/wiki/Sensitivity_and_specificity

[26]This is sometimes called the *specificity* or the *true negative rate*
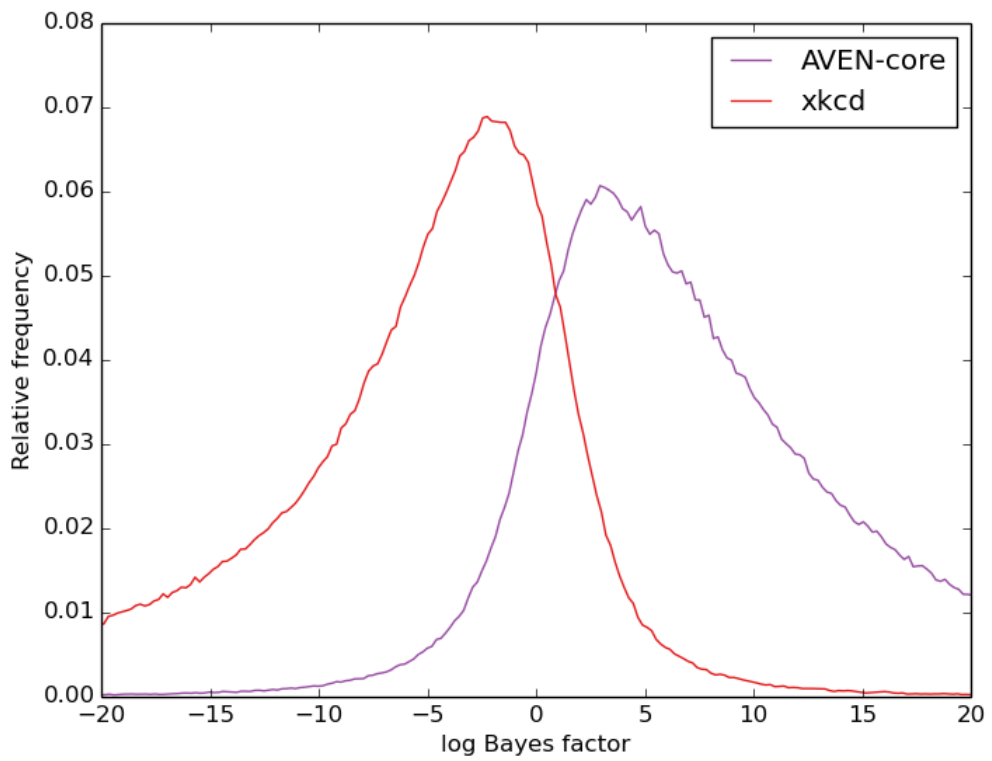
Figure 5.2: The distribution of log_{10} Bayes factor for documents in the testing data for ten trials of the NB classifier.

Forum	% Classified as asexual
Romantic and Aromantic Orientations	97.96 (0.25)
The Gray Area	97.89 (0.27)
Welcome Lounge	97.47 (0.1)
For Sexual Partners, Friends, and Allies	97.05 (0.19)
Asexual Relationships	96.07 (0.12)
Asexual Q&A	94.97 (0.14)
Asexual Musings and Rantings	88.87 (0.14)
World Watch	82.86 (0.66)
Census Forum	82.64 (0.46)
Gender Discussion	81.41 (0.35)
Visibility and Education Projects	78.97 (0.73)
Announcements	76.08 (0.64)
Older Asexuals	75.65 (0.46)
Celebration Time	75.29 (0.58)
Admod Archives	74.17 (0.69)
Site Comments	66.16 (0.62)

Table 5.6: Mean (SD) accuracy for 160 trials of the NB classifier with names for users and AVEN-specific terms removed

these at levels far above chance. However, true positive rates and false positive rates are far below 100%. It is likely that significant improvement would be possible through making use of more information and/or a better approach to smoothing. However, from reading through some of the misclassified documents, I doubt that even human annotators could reach anywhere near 100% accuracy. Not every post in the asexual data is obviously about asexuality or a closely related topic, and there are threads in the xkcd data that are about topics much more frequently discussed on AVEN. Indeed, there is at least one thread in control data that talks about asexuality and even mentions AVEN.

The accuracy of the classifier for the core AVEN data varies considerable by forum. For each forum in the core AVEN data, Table 5.6 shows the true positive rate (and standard deviation) of the classifier for the 160 trials. Depending on the forum, the accuracy ranges from around only one document in fifty being misclassified to around one document in three being misclassified. However, the accuracy for each forum was remarkable consistent across trials–in all cases, SD < 0.75 percentage points.

It is interesting to note that many of the forums with the highest accuracy rate are primarily for people or identities that are, in some sense, at the periphery of asexuality or

asexual communities. The Gray Area is specifically about people/feelings/identities that are neither clearly asexual nor non-asexual. Welcome Lounge is for new members (i.e. not established members of the community). Sexual Partners, Friends, and Allies is primarily for non-asexuals. It may be that comparatively large amount of asexuality-related terminology being used in these forums.

Although the classifier is able to distinguish between the documents labeled "asexual" and "non-asexual" at rates well-above chance, it is possible that the classifier is distinguishing not between asexual discourse vs. non-asexual discourse, but between AVEN discourse and xkcd discourse. Because there is little to believe that non-asexual discourse has anything distinctive about it[27], the relatively high accuracy of classifying xdcd documents (without giving it a high prior probability) and the left-skew in Figure 5.2 suggests that it is xkcd discourse (or possibly some superset it is characteristic of) driving that accuracy. More crucially for my research questions is whether the classifying is distinguishing between xkcd discourse and asexual discourse, or between xkcd discourse and AVEN discourse.

Two pieces of evidence so far suggest that the classifier is distinguishing between xkcd discourse vs. asexual discourse. First, the classifier still performs reasonably well after removing words identified as names for users or AVEN-specific terms. Second, the forums where the classifier had the worst accuracy (Site Comments and Admod Archives) were the ones largely about AVEN itself. (These were also the forums where removing vs. retaining names names for users and AVEN-specific terms had the largest impact on the classifier's accuracy. See Appendix A for details.)

5.2.5 Generalizing the Naïve Bayes classifier

Although the above results suggest that the classifier is distinguishing between *asexual discourse* vs. *xkcd discourse* and not merely *AVEN discourse* vs. *xkcd discourse*, it remains inconclusive. To further investigate the matter, I applied the classifier to two kinds of data:

---

[27]This is probably true in many domains for categories defined in terms of non-X, where X is also a very small part of the the larger domain. e.g. Within the domain "animals", there are various attributes that enable people to identify an animal as belonging to the category "squirrel," whereas there are probably no positive attributes that are characteristic of "non-squirrels".

Forum	% Classified as asexual
Worldpride 2012	62.48 (0.3)
Open Mic	62.0 (0.11)
Off-A	53.24 (0.08)
Just For Fun	45.28 (0.11)
Philosophy Politics and Science	28.17 (0.04)
LJAsexuality	92.95 (0.03)
LJAsexuals	91.9 (0.18)
HHA	85.33 (0.19)
Apositive	85.29 (0.1)
Blogs	82.07 (0.07)

Table 5.7: Mean (STD) accuracy for 160 trials of the classifier (with names for users and AVEN-specifc terms removed) applied to the forums in the AVEN-other corpus and to the communities in the Asexual-other corpus.

- AVEN discourse that is generally not asexual discourse (the AVEN-other corpus)

- Asexual discourse that is clearly not AVEN discourse (the Asexual-other corpus)

Presumably, the classifier's true positive rate being well above chance derives from one of three possibilities: 1) It is detecting asexual discourse. 2) It is detecting AVEN discourse. 3) It is detecting both AVEN discourse and asexual discourse. These make distinct predictions for the classifiers' performance on the non-AVEN asexual data and on the peripheral asexual data.

1. It is detecting asexual discourse rather than AVEN discourse: The true positive rate for the non-AVEN asexual forums should be far above chance. The "true positive rate" for the peripheral AVEN forums should be either around chance or well-below chance.

2. It is detecting AVEN discourse rather than asexual discourse: The true positive rate for the non-AVEN asexual forums should be either around chance or well-below chance. The "true positive rate" for the peripheral AVEN forums should be well-above chance.

3. It is detecting both asexual discourse and AVEN discourse: The true positive rate for both the peripheral AVEN forums and for the non-AVEN asexual data should be well-above chance.

In the same 160 trials as described above, I also applied the classifier to the documents in the peripheral AVEN forums (listed in Table 5.2), and likewise for the non-AVEN asexual communities (listed in Table 5.3). Under the assumption these are all asexual documents, Table 5.7 shows the true positive rate for each forum/community.
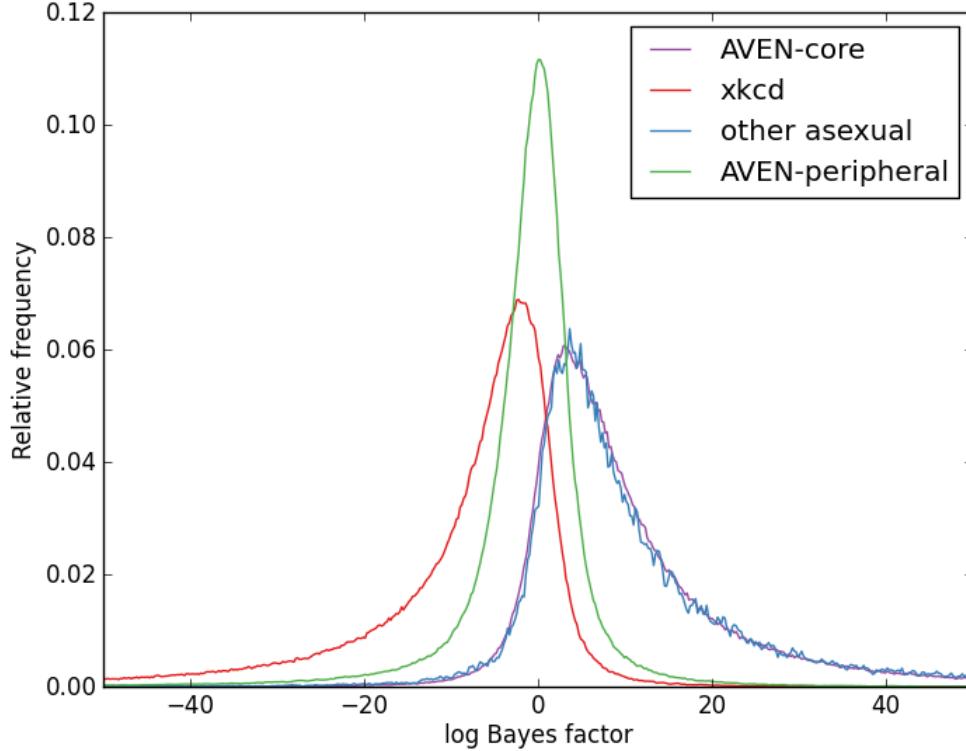
Figure 5.3: Distribution of log Bayes factors for the core AVEN, peripheral AVEN, non-AVEN asexual, and xkcd data in ten trials of the NB classifier.

The results are most consistent with the 2nd hypothesis above: It is (at least primarily) detecting asexual discourse rather than merely AVEN discourse. For all five of the non-AVEN asexual communities, the accuracy (in any condition) ranges from 82.07% to 92.95%. By contrast, the "accuracy" for the peripheral AVEN forums ranged from 28.17% to 62.48%.[28]

Examination of the distribution of log Bayes factors for ten trials (with names for users removed) is shown in Figure 5.3. The distribution of log Bayes factors for the core AVEN data and the non-AVEN asexual data are remarkably similar, while the peripheral AVEN data is quite different from either.

Having established that the classifier can distinguish between asexual vs. non-asexual discourse at levels far above chance, I now want to examine what words are especially strong

---

[28]One could object to this analysis on the grounds that, for two forums, accuracy for the classifier was part of the decision of whether to include it in the AVEN-core corpus or the AVEN-other corpus.

predictors of asexual discourse. Using the words in this way rests on the assumption that such words can (often) be seen as characterizing asexual discourse, and this will be the topic of the next section.

5.3 Keywords

Analyzing lists of "keywords" is a popular method in corpus linguistics (Baker, 2004). A good definition for the meaning of "keywords", as used in corpus linguistics, comes from the online manual for WordSmith Tools, a commonly used corpus linguistic software package: "Key words are those whose frequency is unusually high in comparison with some norm [and which] provide a useful way to characterise a text or a genre"(Scott, 2014).

In producing a workable operationalization to apply to some specific text or set of texts, it is necessary to have some sort of metric of "keyness" as well as a reference corpus. This metric is then calculated on all words (or some subset of these) in the data. Words exceeding some researcher-defined cut-offs are (potentially) considered to be keywords. There are not any widely agreed on standards for acceptable cutoffs, but they are important to a) limit false positive results, b) restrict the list to words with reasonably strong associations, and c) keep the size of the list manageable (given that subsequent analysis is largely qualitative). Frequently, there will be a few additional requirements to weed out certain sorts of words: For example, words that only occur a small number of times in the study corpus, or words that are highly concentrated in one part (or a small number of parts) of the corpus.

For my keyword analysis, the method follows straightforwardly from the NB classifier, with a few modifications. The study corpus will be the core AVEN data and the reference corpus will be the xkcd control (only there is no need to set aside testing data). The keyness metric will be closely related to the Bayes factors used above. As Kilgarriff (2009, 2012) observes, the smoothing value can be adjusted based on how much the researcher wants to bias the model towards higher or lower frequency words. He proposes a smoothing value of

100 as a default, and I will follow him in this regard.[29]

From examination of a list of keywords generated by this method, it seems to work fairly well, but there are two main kinds of words that can receive high keyness scores that we want to remove: Names for users and words extremely concentrated in a small part of the corpus. Names for users have been removed with the same method described above in Section 5.2.3, although AVEN-specific terms have been retained.

It is well-known in corpus linguistics that a rare word with high frequency in just one or two documents sometimes appear in keyword lists, and various methods have been proposed for dealing with this. The main kinds of concentration-in-a-single-part-of-the-corpus that need to be removed are words overwhelmingly concentrated in a small number of threads, or in posts of a small number of users.

Using a measure for dissemination of words across users [or threads] proposed by Altmann et al. (2011), for each word occurring at least 5 times, I took the ratio of the observed number of users who used it and the expected number of users (under a bag of words model in which the word-count for each word in the corpus and each user are both held constant). And likewise for threads. Words were removed from the keyword list if their dispersion value was in the bottom 1% of words, when grouped by threads and/or by users. (For discussion of why most of the metrics more commonly used in corpus linguistics are poorly suited for my data, as well as the details of these calculations, see Appendix B).

Using this method, I calculated the keyness metric for each word used at least once in the AVEN-core corpus (other than the words excluded because of the above), sorted these, and selected the top 200 for subsequent analysis. This number was chosen as it is large enough to give a good sense of the major terms, topics, and themes of asexual discourse, while being small enough to be manageable (after lemmatizing, the number is around half this).

---

[29]Kilgarriff's keyness metric is slightly different than the one I will use because he does not adjust the total words in the corpus to take into account the words added via smoothing.

5.3.1 Results

Of the 200 word-forms identified by the keyword analysis, all appear at least 519 times in the study corpus (about 7 times per million words). All but 6 also appear in the xkcd data at least once, and of those appearing in both datasets, all occur at least 5.5 times as often in the AVEN data as in the xkcd data. As Kilgarriff (2012) observes, when word-forms are used (rather than lemmas), a single lemma will often appear multiple times: Although redundant, this redundancy provides additional evidence of the importance of the lemma. The top 30 words, ranked by keyness value, are shown in Table 5.8, along with their use per million words in the AVEN-core corpus, in the control corpus, and the keyness value. The full list may be found in Appendix C.

As is common practice with keyword analysis, I have grouped the terms for purposes of presentation. Most categories are thematic, but some are defined in linguistic terms. No inferential statistical analyses rely on which category a word is placed in, so reliability is less important than highlighting recurrent themes. In some cases, a word could have plausibly fit into multiple categories. In these cases, I used the following heuristics:

1. If a word fits into a thematic category and a linguistic category, it is included in the linguistic category.

2. When possible, words sharing a root are grouped together.

Among the words, the roots sex and *romance/romantic* were particularly prominent, and so the first four categories are primarily linguistically defined, based on these: words containing or derived from *asexual* (possibly metaphorically), other words containing the root *sex*, words containing the root *romance/romantic*, and prefixes clipped from words containing *sexual* (other than *a-*). Readers unfamiliar with some of the terms are referred to the AVENwiki.[30]

As can be seen, there are numerous terms referring to types of people (categorized in ways relating to their sexuality), attitudes about sex (e.g. *anti(-)sexual*, *sex-positive*, *sex-repulsed*),

---

[30]http://asexuality.org/wiki/index.php?title=Lexicon

	keyword	asex rfreq	ref rfreq	keyness value
1	aven	1414.97	0.6	2.85
2	asexuality	1284.8	2.44	2.53
3	asexual	2721.17	7.46	2.5
4	asexuals	684.79	1.29	2.41
5	sexuals	308.91	0.23	2.28
6	aromantic	262.75	0.27	2.2
7	demisexual	85.96	0.16	1.75
8	attraction	934.58	15.71	1.75
9	ace	284.28	4.11	1.72
10	repulsed	128.35	1.21	1.7
11	romantic	874.11	18.47	1.65
12	libido	151.21	2.77	1.57
13	avenites	49.84	0.03	1.55
14	asexy	51.32	0.1	1.55
15	intimacy	143.83	2.85	1.54
16	romantically	105.54	1.8	1.53
17	aces	74.72	0.93	1.52
18	arousal	109.62	2.27	1.49
19	omg	41.41	0.09	1.46
20	welcome	1390.56	48.58	1.46
21	aroused	114.27	2.76	1.45
22	crushes	123.89	3.12	1.45
23	grey-a	38.11	0.0	1.45
24	cuddling	132.6	3.48	1.45
25	demi	49.04	0.46	1.44
26	kissing	249.02	8.02	1.43
27	heteroromantic	37.68	0.09	1.42
28	celibate	66.09	1.42	1.38
29	sexual	3019.27	134.79	1.36
30	gray-a	30.1	0.03	1.34

Table 5.8: Top 30 keywords for AVEN-core (as compared against)

Category	Num.	Keywords
asexual	11	ace, aces, asexual, asexuality, asexually, asexuals, asexy, gray-a, grey-a, non-asexual, amoeba
sex (excluding asexual)	23	anti-sexual, antisexual, autosexual, bisexual, demi-sexual, demi-sexual, demisexuality, demisexuals, heterosexual, hypersexual, hyposexual, non-sexual, nonsexual, pansexual, sex, sex-positive, sex-repulsed, sexless, sexual, sexualities, sexuality, sexually, sexuals
romance/romantic	18	aro, aromantic, aromanticism, aromantics, bi-romantic, biromantic, demiromantic, hetero-romantic, heteromantic, heteroromantic, homo-romantic, homoromantic, non-romantic, panromantic, romance, romantic, romantically, romantics
clippings	4	bi, demi, demis, hetero

Table 5.9: Keywords containing or derived from words with the root *sex* or *romance*, and the number of word-forms per category

some informal terms (e.g. *amoeba*, *asexy*), and one out-group terms (e.g. *non-asexual*). There are several terms relating romance and romantic attraction, including romantic orientation terms, and terms reflecting a distinction between people who are aromantic and those who are not–in particular, the inclusion of *romantics* (a plural noun) suggests that many of the uses of romantic also reflect a nominal usage. Noticeably absent from this list is *allosexual*, which has become the primary out-group term in tumblr asexual discourse.[31]

The list also contains several intermediate category terms, including hyposexual and two spellings of gray-a. The prefix *demi-* is particularly prominent, with four forms of *demisexual*, two of the clipped *demi*, and also *demiromantic*, which is case of the prefix being extended into the romantic-orientation terminology. These provide further evidence for the importance of intermediate category terms in asexual discourse.

Among the keywords that did fit into the above categories, several themes emerged and I have grouped words accordingly:

1. sex and/or sexuality (including words that often–but not always–have a meaning relating to sex/sexuality),

2. attraction, emotions, and (un)desire

3. relationships

4. gender (especially gender identity)

5. interpersonal contact behaviors that may or may not be sexual

6. websites and website-related terms

7. greetings

8. miscellaneous.

---

[31] https://asexualagenda.wordpress.com/2012/09/12/why-i-use-allosexual/

Category	Num.	Keywords
sex and/or sexuality	34	arousal, aroused, arousing, celibacy, celibate, erection, fantasies, fantasize, fantasizing, fetish, fetishes, foreplay, frigid, horny, intercourse, libido, masturbate, masturbated, masturbating, masturbation, nonlibidoist, orgasm, orientation, orientations, prude, queer, lesbian, lgbtq, repressed, stimulation, tmi, virgin, virginity, virgins
attraction, emotions, and (un)desire	32	aesthetic, aesthetically, affection, affectionate, attracted, attraction, attractions, attractive, crave, crush, crushes, desire, desiring, disgusted, disinterest, emotionally, feelings, felt, grossed, grosses, indifferent, pleasing, pleasurable, relieved, repulsed, repulses, repulsion, repulsive, sensual, squish, squishes, urges
relationships	26	boyfriend, boyfriends, closeness, companionship, dated, dating, flirt, flirting, friendship, friendships, husband, intimacy, intimate, loneliness, lonely, loner, partner, partner's, partnered, partners, personals, platonic, relate, relationship, relationships, soulmate
gender	10	agender, agendered, androgyne, androgynous, androgyny, ftm, genderless, genderqueer, neutrois, tomboy
behaviors	8	cuddle, cuddles, cuddling, hugging, kiss, kissed, kisses, kissing
greetings	7	cake, greetings, hello, hi, welcome, welcomes, welcoming
websites	13	acebook, aven, aven's, avenite, avenites, avenwiki, chat, faq, faqs, musings, off-a, rantings, tumblr
miscellaneous	14	aspergers, aspie, bloomer, congrats, introvert, introverted, label, labels, montel, omg, schizoid, sherlock, therapist, visibility

Table 5.10: Other keywords for asexual discourse, and the number of word-forms per category

Not surprisingly, asexual terminology, along with sex and relationships, feature prominently in these words. The inclusion of several words pertaining to gender identity and attributes—especially terms for non-binary gender identities (e.g. *agender*, *genderqueer*, *neutrois*)—is consistent with the finding that around 10 - 20% of participants (or at least survey-takers) in English-language online asexual discourse identify as neither male nor female.[32]

The words about attraction and emotion strongly suggest that both desire (e.g. *crave*, *desire*, *urges*) and undesire (e.g. *disgusted*, *grossed*, *grosses*, *repulsed*, *repulses*, *repulsion*, *sex-repulsed*) are prominent features of asexual discourse, as is non-desire (e.g. *disinterest*, *indifferent*).

---

[32]It seems that this does not for online Spanish asexual communities, according to annual surveys conducted by Spanish AVEN from 2011: http://asexuality.org/sp/encuestas

word	frequency
hyposexual	487
semisexual	102
gray-a	31
demisexual	24
pseudosexual	22

Table 5.11: Frequency of intermediate category terms on AVEN at the end of 2006.

5.4 Usage over time

The asexual corpora are especially well-suited for examining word usage over time. This section will focus on the development of three domains of asexual-related terminology: intermediate category terms, romantic-attraction terms, and forms of *repulsed* and *indifferent.*

5.4.1 Intermediate categories

Chapter 4.5 discussed qualitative evidence for intermediate-category terms, and highlighted a thread from October 2006 about the various intermediate-category terms that gives a good idea of the state of asexual terminology for intermediate categories at that time. Four terms were listed *hyposexual, semisexual, pseudosexual,* and *demisexual,* suggesting that *hyposexual* was the most common term. Another term (*gray-A*) was suggested. Table 5.11 shows the frequency of these five terms in the AVEN data at the end of 2006, which confirms that *hyposexual* was the most commonly used.[33]

The use of *demisexual, hyposexual,* and *gray-a* is shown in Figure 5.4. For the first several years, *hyposexual* was used moderately often, but its use declined considerably once *gray-A* came into prominence. Although *demisexual* was coined somewhat before *gray-A*, it came into more prominent use shortly after *gray-A* did. Since then, it has become used even more. (The difference would be more obvious if the clipping *demi* was included.) These terms have become much more frequent than *hyposexual* ever was. The graph demonstrates that there has been a rise of intermediate categories in asexual discourse, and that the most commonly

---

[33]Because *hypoactive sexual desire disorder* was sometimes called *hyposexual disorder*, uses of *hyposexual* were excluded if it was followed by *disorder* or *desire disorder*. Prior to the thread mentioned in the text, *gray-A* had only been discussed in one thread. This thread is not in the AVEN corpus, presumably because of the rare issue mentioned in 7, so a manual count was added to this table.

used terms started to come into prominence beginning in late 2007 – early 2008.

5.4.2 Romantic Orientation

Another noteworthy shift in asexual discourse involves terminology concerning romantic orientation. In Chapter 4.5.1, I argued that although romantic-orientation terms have been in used in online asexual discourse since before the creation of AVEN's forums, their prominence has increased, especially in introductory materials about asexuality.

In the first several years represented in the AVEN-core and Asexual-other corpora, the most common romantic-orientation-related terms were *hetero-asexual*, *homo-asexual*, *bi-asexual*. Since then, *biromantic*, *heteroromantic*, and *homoromantic* have become more common: All three of these (with and without the hyphen) appear in Table 5.9, and for *hetero-* a third form also appears (*heteromantic*).

Figure 5.5 shows the usage (per million words) over time for the main forms of *X-asexual* vs. *X-romantic*, when restricted to the prefixes *hetero-*, *homo-*, and *bi-*. As can be seen, the *X-asexual* terms were more commonly used than the *X-romantic* ones until about 2006. For a period in 2004 and early 2005, the use *X-asexual* and *X-romantic* both increased, and then *X-asexual* started to fall, while *X-romantic* continued to rise for several years.

For the first few years of their use, hyphenated and unhyphenated versions of *X-romantic* were both quite common, as seen in Figure 5.6. Then around 2009 or 2010 the unhyphenated versions continued to rise, while the hyphenated versions were used less (although both forms remained common). This suggests that the words had been an increasingly established part of asexual discourse.

A plausible explanation for why *X-romantic* replaced *X-asexual* involves increased discussion about *aromanticism*. When comparing the two paradigms, it quickly becomes apparent that *X-romantic* has plausible parallels to *asexual*, *bisexual*, *heterosexual*, and *homosexual*, while *X-sexual* is only easily adaptable for the last three of these, because a-asexual involves a reduplication the prefix *a-*. To illustrate, in the *X-romantic* paradigm, *aromantic* occurs more often than the *bi- hetero-*, and *homo-* forms combined (See Figure 5.7). By contrast,
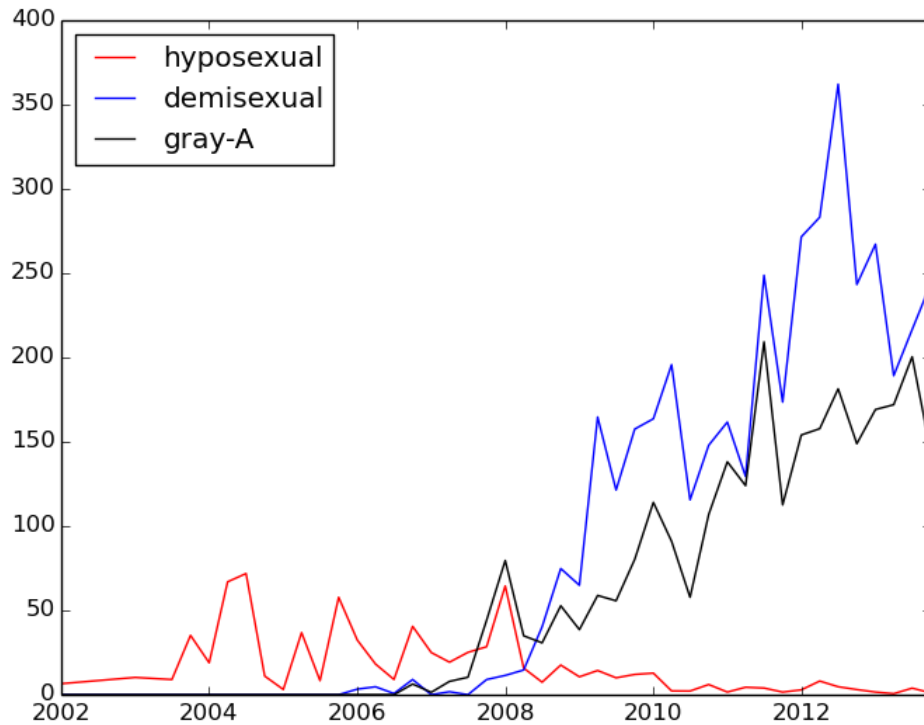
Figure 5.4: Usage (occurrences per million words) over time for *gray-A*, *demisexual*, and *hyposexual* in the AVEN-core and asexual-other corpora. Language is divided into 6-month segments, except for the first time period (due to sparsity of data) which covers 2000–2002. Counts include hyphenated and unhyphenated forms; both *grey* and *gray* were included, along with *gray-asexual* and *gray-sexual*. In addition, for *sexual*, words were counted if they ended in *-s*, *'s*, or *ity*. Uses of *hyposexual* were excluded if followed directly by *disorder* or *desire disorder*. The regex used are as follows (word-breaks are omitted here, as are the exclusionary criteria for *hyposexual*):

hypo-?sexual('?s|ity)?
gr[ea]y-?(a|a?sexual)('?s|ity)?
demi-?sexual('?s|ity)?
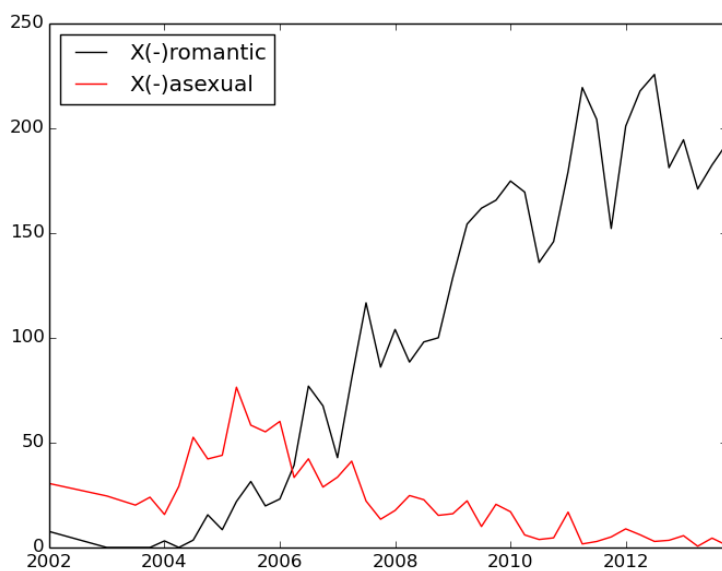
Figure 5.5: Usage (occurrences per million words) for *hetero-*, *bi-*, *homo-asexual* vs. *hetero-*, *bi-*, *homo-romantic* in the AVEN-core and asexual-other corpora. Time is measured in 3-month segments, except for the first time period (due to sparsity of data) which covers 2000–2002. Counts for *X-romantic* also include words ending in *s*, *'s*, or *ism*. Counts for *X-asexual* also include words ending in *s*, *'s*, or *ity*. The regex used are as follows (word-breaks are omitted here):

(hetero|homo|bi)-?romantic('?s|ism)?
(hetero|homo|bi)-?asexual('?s|ity)?
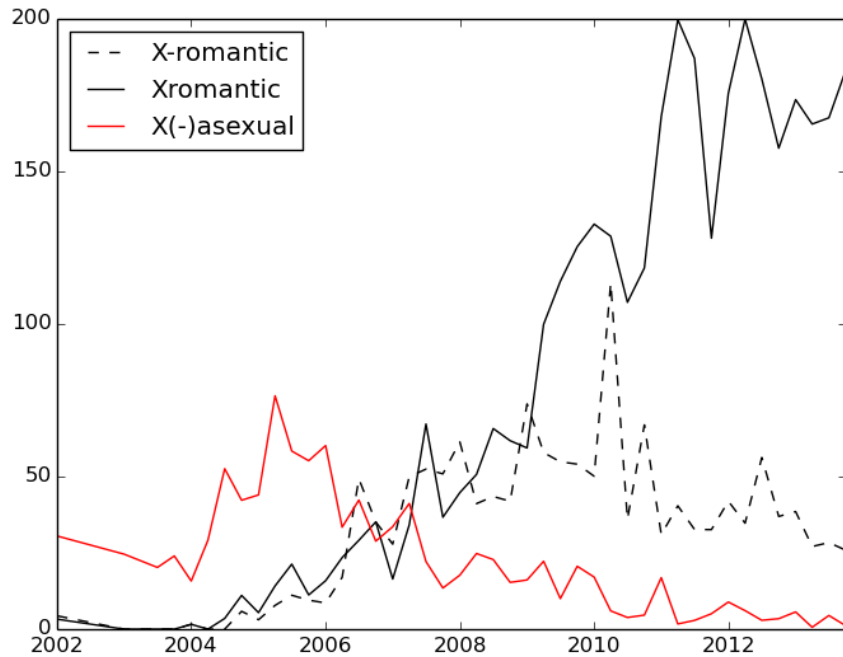
Figure 5.6: Usage (occurrences per million words) for hyphenated vs. unhyphenated forms of *bi/hetero/homo(-)romantic* in the AVEN-core and asexual-other corpora, shown alongside *bi/hetero/homo-asexual* (regardless of hyphenation). Time is measured in 3-month segments, except for the first time period, which covers 2000–2002. For the *romantic* words, tokens ending in *-s*, *'s*, and *ism* were also counted. For the *sexual* words, The regex used are as follows (word-breaks are omitted here):

(hetero|homo|bi)-(ro)?mantic('?s|ism)?
(hetero|homo|bi)(ro)?mantic('?s|ism)?
(hetero|bi|homo)-?asexual('?s|ity)?
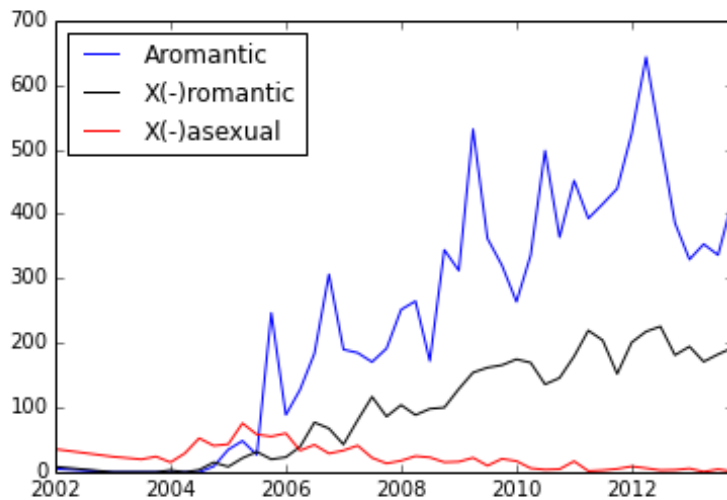
Figure 5.7: Usage (occurrences per million words) for *aromantic*, along with *bi-*, *hetero-*, and *homoromantic* (combined as *X-romantic*) and *bi-*, *hetero-*, and *homo-asexual* (combined as *X-asexual*) in the AVEN-core and asexual-other corpora. Time is measured in 3-month segments, except for the first time period (due to sparsity of data) which covers 2000–2002. Counts for *X-romantic* also include words ending in *s*, *'s*, or *ism*. Counts for *X-asexual* also include words ending in *s*, *'s*, or *ity*. Counts include both hyphenated and unhyphenated forms. The regex used are as follows (word-breaks are omitted here):
a-?romantic('?s|ism)?
(hetero|homo|bi)-?romantic('?s|ism)?
(hetero|homo|bi)-?asexual('?s|ity)?

forms of *a-asexual* are attested only 73 times, as compared with 350 times for *homo-asexual*, 457 for *hetero-asexual*, and 625 for *bi-asexual*. Of these, some are typos (e.g. some uses of *aasexual*), some mean "not asexual", and a few mean "very asexual".

Figure 5.7 shows the relative usage over time for the same *X-asexual* and *X-romantic* words as above, along side usage of *aromantic*. Use of *aromantic* starts to rise around the same time as the other three *X-romantic* forms. Since around 2005, *aromantic* has been used more than the other three (combined), and it was around this time that *bi-, hetero-, homo-romantic* began to replace *bi-, hetero-, homo-asexual*, suggesting that *aromantic* played a leading role in the *X-romantic* paradigm replacing the *X-asexual* paradigm.
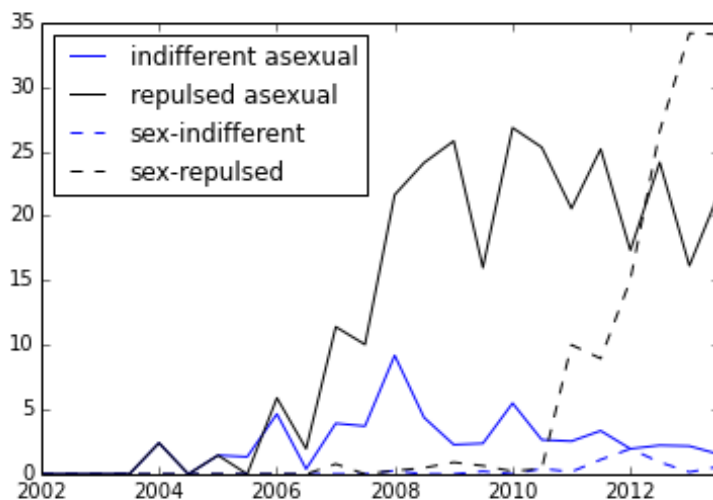
Figure 5.8: Usage (occurrences per million words) for *sex-repulsed*, *sex-indifferent*, *repulsed asexual*, and *indifferent asexual* in the AVEN-core and asexual-other corpora. Time is measured in 6-month segments, except for the first time period (due to sparsity of data) which covers 2000–2002. Counts for *X-asexual* also include words ending in *s*, *'s*, or *ity*, and also include *X-ace*. Counts include separating these with a hyphen of a space. The regex used are as follows (word-breaks are omitted here):

indifferent[-]a(ce|sexual)('?s|ity)?
repulss?ed[-]a(ce|sexual)('?s|ity)?
sex-indifferent
sex-repulss?ed

5.4.3 Repulsed and Indifferent

In Chapter 4.5.4, I discussed the terms *indifferent* and *repulsed* in asexual discourse, looking at the AVENwiki and four non-AVEN glossaries of asexuality-related terminology. These suggested that a variety of forms are used, including *repulsed*, *indifferent*, *repulsed asexual*, *indifferent asexual*, *sex-repulsed*, and *sex-indifferent*. The evidence suggested that *repulsed* is more prominent than *indifferent*. Further evidence for this comes from these (and closely related) words in keywords in Tables 5.9 and 5.10: *indifferent*, *repulsed*, *repulses*, *repulsion*, *repulsive*, and *sex-repulsed*. Both *repulsed* and *indifferent* appear in this list, but there are five forms of *repulsed* and only one of *indifferent*.

Usage over time for *Y(-)asexual* and *sex-Y* (where *Y* is *repulsed* or *indifferent*) is displayed in Figure 5.8. As can be seen, none of these forms were attested in any of the earliest time

106

periods. From 2004–2006, both forms of *Y-asexual* were used at similar (though not very high) rates. From 2007-2008, the usage of *repulsed asexual* increased dramatically, and has remained at similar levels since then. By contrast, *indifferent asexual* continues to be used, but much less than *repulsed asexual*. The *sex-Y* forms were first attested in 2007, but remained rare until 2011. Since then, *sex-repulsed* increased dramatically, surpassing *repulsed asexual*. However, *sex-indifferent* has not gained much currency.

The fairly restricted forms shown in Figure 5.8 give us reasonable confidence that these terms are specifically about types of asexuals and/or feelings about sex. However the forms *repulsed* and *indifferent* are much more common than any of the terms shown there. Usage for these is displayed in Figure 5.9. Because these terms alone give little idea of indifference or repulsion to what, the graph also displays the usage of these terms where they are followed by a word beginning with *sex* in the next five words.

As can be seen, the terms *indifferent* and *repulsed* are used at roughly comparable rates through early 2007, but afterwards *repulsed* came to be used considerably more than *indifferent*. This pattern reflects the trends for *Y asexual*, suggesting that the increase in usage of *repulsed* reflects the increased use of *repulsed asexual*.[34] For both *repulsed* and *indifferent*, a sizable minority of these occur shortly before a word beginning with *sex*, which is consistent with the idea that a large proportion of these uses relate to feelings concerning sex.

Evidence from usage over time, multiple lists of asexuality-related terminology, and keywords lists all suggest that both *undesire* and *nondesire* are important parts of asexual discourse. Because of the lexical focus on only two sets of terms for expressing these, many others are missed. Still, some trends stand out. First, for this pair of terms, *repulsed* has become the much more prominent of the two. Second, this difference is much more pronounced for *sex-Y* than it is for *Y asexual*.

The reasons for this difference are unclear. One plausible explanation is that it reflects

---

[34]Of course, uses of *repulsed asexual* are a subset of uses of *repulsed*. However, this alone does not account for the increased usage of *repulsed*. In 2003, usage per millions words for *repulsed* was about 35; in 2013, it was about 221. In this same time period, *repulsed asexual* increased from 0 to 19. Furthermore, when looking at one-year periods, *repulsed asexual* peaked at about 26 in 2010, and decreased by about 2 each subsequent year. By contrast, *repulsed* saw a 71% increase during the same period.
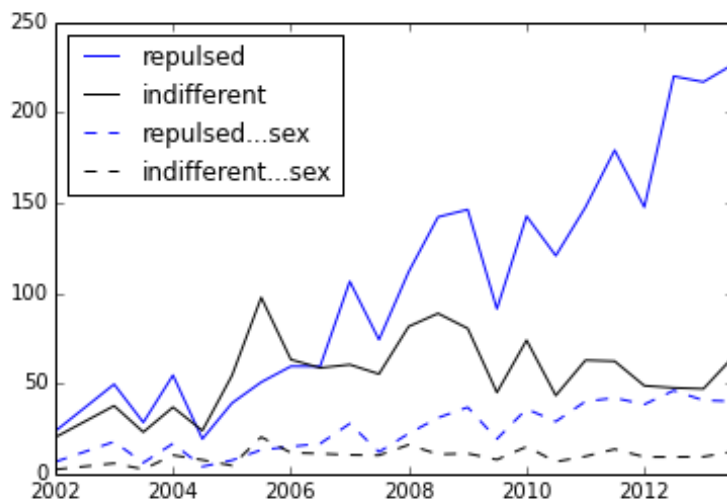
Figure 5.9: Usage (occurrences per million words) for *repulsed* and *indifferent* in the AVEN-core and asexual-other corpora, along with usage per million words for these two terms, where a word starting with *sex* occurs within the next five words. Time is measured in 6-month segments, except for the first time period (due to sparsity of data) which covers 2000–2002. The regex used are as follows (word-breaks are omitted here):

indiffe?rent
repulss?ed
indiffe?rent ([a-z]+)0,4sex[a-z-]*
repulss?ed ([a-z]+)0,4sex[a-z-]*

attitudes among participants in asexual communities, but the best available evidence suggests that, at most, this accounts for a small portion of the difference. The 2014 AVEN Community Census (Ginoza et al., 2014) asked people "to rate how they feel towards the idea of themselves engaging in sexual activity, picking from the labels of 'repulsed', 'indifferent', and 'favorable'." Among asexuals there is about a 1.3:1 ratio of repulsed to indifferent. Among gray-asexuals the ratio was about 1:2.3, and for demisexuals about 1:1.9. Further research is needed to understand why this is the case.

5.5 Conclusion

In this chapter I have discussed the design of multiple corpora of online asexual discourse, as well as a reference corpus. The first research question was whether there are characteristics

of online asexual discourse that distinguish it from other online discourse. Through the use of a Naïve Bayes classifier, I found evidence that there are distinguishing features. To ensure that differences found were truly characteristic of asexual discourse—and not merely AVEN discourse—I used several methods to remove names for users and AVEN-specific terms. I ran the classifier on AVEN non-asexual discourse and on non-AVEN asexual discourse and found that the non-AVEN asexual data patterned with the AVEN asexual data, while the AVEN non-asexual data did not. I then performed an analysis of the top 200 keywords, finding that (a)sexuality, relationships, and attraction and emotions feature prominently. I found evidence that desire, undesire, and non-desire are all prominent features of asexual discourse. I then performed follow-up analyses on three domains found in the keywords: intermediate category terms, romantic orientation terms, and the terms *repulsed* and *indifferent*. I found that there have been considerable changes over time for each of these domains.

Chapter 6

Survey about sexual identity before and after finding online communities

In the last chapter, I investigated online asexual discourse through use of corpora of naturally occurring language data. This enabled us to examine the content of asexual discourse, but such data cannot tell us about how encountering online asexual discourse affected asexuals' understanding of asexuality, and it provided limited opportunity to test the *AVEN shift* hypothesis.

In this chapter, I report on the findings of a survey conducted online. The purpose of the survey was to investigate the hypothesis that there has been an *AVEN shift*. The *AVEN shift* hypothesis involves changes in people's understanding and self-application of terms that have been in use since the earliest attested online asexual discourse, and so one of the only possible ways to investigate changes from before the emergence of asexual communities is via retrospective reporting.

Because one major purpose was to investigate how asexuals growing up in English-language contexts understood their asexuality prior to the rise of online asexual communities, participation was restricted to asexual-identified individuals 25 years of age or older (as of January 2012) who had lived most of their lives since age 12 in English language contexts. Age 12 was used as a proxy for puberty, and 25 was used to approximate this occurring before the year 2000, which roughly corresponds to when there began to be considerable growth in online asexual communities, as seen in Chapter 4. The language requirement was because the survey had retrospective questions about a number of English words; as such being a native speaker of English was unnecessary, but only requiring present proficiency in English would not be strict enough.

6.1 Introduction

The goal of the survey was to investigate asexuals' understanding of asexuality prior to the rise of online asexual communities (or at least prior to finding an online asexual community). The survey investigated conceptual sources for understanding asexuality prior to the creation of online asexual communities, and it also investigated the AVEN shift. In particular, it tests five hypotheses predicted by the hypothesis that there has been an AVEN shift:

1. There should be a decrease in the number of people identifying as *nonsexual*.
2. There should be a decrease in people considering themselves *abstinent* or *celibate*.
3. People's understanding of asexuality should move in the direction of "not experiencing sexual attraction."
4. People's understandings of asexuality should move away from "not sexual" or "lacking sexuality."
5. People's understanding of asexuality should move in the direction of "sexual orientation" and seeing themselves aligned with LGBT.

The fourth of these hypotheses relates to an understanding of asexuality described in a 2006 lecture by David Jay which is linked to on the AVENwiki:

> [Asexuality is] more about disidentifying with sexuality [than sharing a common trait]. So the real common experience within the asexual community isn't just that we all don't experience sexual attraction. It's that all of us have looked at this very complicated thing in the world called "sexuality" and said, "You know what? That is not me. I can't relate to it. I don't wanna deal with it. I wanna find a way to live my life without it."

A similar understanding is expressed in much of the academic literature about asexuality, especially from the humanities and the exclusively qualitative social sciences—Some authors will interpret *asexual* (as attributed to people) as meaning "without sexuality." For example, Scherrer (2008, p.360) calls asexuality an "identity [that] revolves around the lack of sexuality," Kim (2011, p.488) uses the phrase "absence of sexuality" as a synonym for asexuality,

and Bishop (in Bishop, 2013, p. 100) states, "Asexuality is identifying oneself as having a lack of sexuality."

In contrast to this, my general observation has been that, among self-identified asexuals, some do consider their asexuality as a lack/absence of sexuality, while some do not. There are at least some people who identify as asexual because they feel the not-experiencing-sexual-attraction definition accurately describes them, although they do not have the "gut feeling" described in the above quote. I have suspected that 1) viewing asexuality as a "lack of sexuality" was more common early on in asexual communities than it is at present, and that, if this is true, 2) part of the reason for this may be that people who, independently, thought of calling themselves *asexual* are more likely to feel that they "lack sexuality" than those who do not. Directly testing the first of these is not possible with my current survey—I would need a much larger number of people who first joined an asexual community before 2004. However, establishing the second would strongly suggest the first because, as Jay (2003, p. 6- 7) observes, as of 2003, "AVEN's access statistics show that most users arrive at the site either by typing the word 'asexuality' into a search engine or by typing directly the address....", and thus the make-up of the community was constrained by having "access to a computer with an internet connection and the technical expertise to use the internet. One must then feel compelled to type either 'asexual' or 'asexuality' into a search engine."

Results generally confirmed all five main hypotheses, but they show that the changes were by no means monolithic. I also found some evidence that people who independently thought of applying the term *asexual* to humans and considered themselves asexual before finding an online asexual community were more likely to regard their asexuality as a lack of sexuality, but the evidence is not very strong.

Because the quantitative survey data points in the same direction as the three lines of qualitative archival data discussed in Chapter 4.5, this adds additional strength to the hypothesis that asexual discourse has shifted in the direction on how AVEN presents asexuality.

6.2 Methodology

The survey was conducted from January 6th to 12th 2012. It was for asexual-identified individuals 25 years of age or older who have lived most of their lives since age 12 in English language contexts. The age requirement was to limit participation to people who were adolescents or older when sizable online asexual communities began to emerge in the early 2000s, and the language requirement was because a number of items asked about specific (English) terms. The survey was hosted on SurveyMonkey (surveymonkey.com), a website that hosts online surveys. No information was collected about what other languages (if any) people grew up using. A copy of the survey can be found in Appendix D. IRB approval can be found in Appendix E.

Links to the survey were posted on AVEN, LJ Asexuality, the AVEN PT Tumblr[1], and Ace-Book [2] after getting permission from those who manage them. They were posted on the blog *Asexual Explorations Blog*[3], and that link was reposted by *Asexual News* [4], a site that frequently posted calls for participants in studies recruiting asexuals. Because SurveyMonkey allows for the use of multiple collectors,[5] one collector was used for AVEN, one for LJ Asexuality, one for Ace-Book, one for the AVEN PT Tumblr, and one for the blog. Because Asexual News reposted the blog link, it is impossible to distinguish these sources of participants. At the time of collection, the AVEN PT tumblr was very new and had few followers, so few people came from Tumblr.

After clicking on the link, participants were taken to the consent form, and those who gave informed consent were then taken to the survey. For each participant, the survey had between 22 and 57 questions, depending upon certain responses, as well a comment box at the end. After giving informed consent, participants were asked about some demographic information, about finding and joining online asexual communities, and some questions about

---

[1]http://avenpt.tumblr.com/

[2]ace-book.net

[3]asexystuff.blogspot.com

[4]Formely on http://www.asexualnews.com. It is archived on the Way Back Machine.

[5]SurveyMonkey allows for multiple "collectors for surveys. Each collector is a different url that links to the survey and then includes which collector someone came from in the dataset.

asexual identity.

Based on this, respondents were asked whether they considered their asexuality to be a lack of sexuality, with the options of *yes*, *no*, and *other*. The reason that "lack of sexuality" was used rather than "not sexual" is that "sexual" has come to be used as an out-group term to mean "not asexual" and it was possible that some respondents would understand the question in this sense if asked whether they consider themselves "not sexual" (p.c. Michael Doré).

For each of the terms *asexual*, *nonsexual*, *celibate*, and *abstinent*, there was a page with a screening question asking whether participants had ever identified as that.[6]If they responded *yes*, they were taken to the following page asking several questions. A *no* for asexual took people to the end of the survey (n=2) and their data was removed from analysis. A *no* for the other 3 words skipped the page asking about that term. Participants were asked about two diagnoses (Hypoactive Sexual Desire Disorder and Sexual Aversion Disorder), each of which had two screening questions (Had they ever heard of it before? Had they heard of it before finding an online asexual community?)[7]

Participants were also asked whether they had ever invented any words to describe what they now consider asexuality, whether they had ever participated in an offline asexual organization (i.e. that used the word *asexual*) prior to finding an online asexual community, and whether they had ever participated in an offline community for issues related to what they now consider asexuality. Following this was a series of questions, first about their relationship to LGBT groups and identity. Some items had an "other: specify" option. For some of these items, some of the responses could be back-coded to one of the established options (in all cases a *yes* or a *no*.) For completed surveys, the median completion time was 11 minutes 20 seconds.

---

[6]Questions 9, 16, 24, and 29 in Appendix D
[7]Questions 34 and 40 in Appendix D
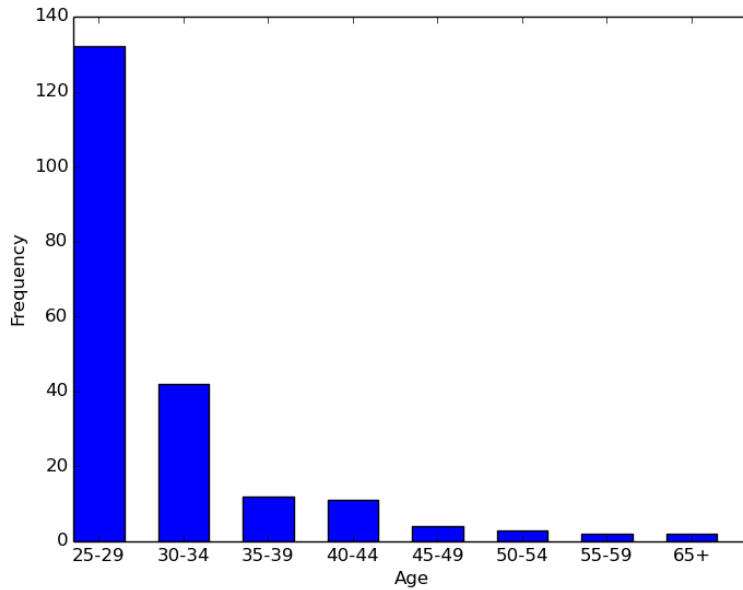
Figure 6.1: Reported age of respondents for the survey.

6.2.1 Participants

The survey was online for four days, after which time I closed it because I had a sufficient number of responses. A total of 289 individuals started the survey, with a total of 208 completed responses. Data for uncompleted surveys was removed from analysis. In the gender item, 146 (70.1%) identified themselves as female, 26 (12.5%) as male, and 36 (17.3%) as something else. Based on the unexpectedly low proportion of "male," responses were examined by collector for collectors receiving at least ten responses (which was all but the AVEN PT Tumblr). The proportion of male responses ranged from 3% (LiveJournal) to 33% (AVEN), which is consistent with the prediction of Hinderliter (2009b, p. 620) that "recruiting [from the LiveJournal Asexuality community], in addition to AVEN, is likely to substantially increase the number of responses but may introduce a gender confound because LJ might have a strong bias towards female users." Age was divided into five year categories (25-29, 30-34 ,.., 60-64, 65+), and results are displayed in Figure 6.1, which shows that respondents overwhelmingly indicated belonging to the youngest age bracket.

6.3 Results

6.3.1 The meaning of specific terms

Participants were asked whether they considered their asexuality to be a "lack of sexuality." After back-coding *other* responses that clearly indicated either a *yes* or *no*, 78 (37.5%) of the respondents indicated that they did consider their asexuality to be lack of sexuality, 119 (57.2%) indicated that they did not, and 11 (5.2%) could not be coded as *yes* or *no*.

Questions for *asexual* and *nonsexual* were mostly analogous to each other. Respondents were asked whether they called themselves *asexual/nonsexual* prior to finding an online asexual community and whether they thought of the term asexual/nonsexual (as something applying to humans) on their own. (For the *asexual* question, respondents were told to mark *yes* if they had learned the term in the context of asexual reproduction and then extended the meaning to describe themselves.) For *nonsexual*, they were asked if they presently consider themselves *nonsexual*.[8] They were asked how they understood the term when they first called themselves asexual/nonsexual, and how they understand the term now. Four options were given for these items and participants were told to check all that apply:

- Not sexual
- On analogy with hetero/homo/bisexual
- Not sexually active
- Other-specify

Table 6.1 shows the results for the questions about identifying as asexual prior to finding an online asexual community and about independently thinking of the term *asexual* as applying to people. A sizable minority of participants (38.8%, 95% CI = [32.15%, 45.52%]) indicated that they had identified as asexual some time prior to finding an online asexual community. Not surprisingly, independently coining/adapting *asexual* was much more common among those who identified as asexual prior to finding an online asexual community

---

[8]There was no corresponding question for *asexual* because of the participation requirements.

(50% did, 95% CI=[38.99%, 61.01%]) than among those who did not identify as asexual prior to finding an online community (13% did, 95% CI=[7.5%, 19.49%]).

		Call self *asexual* pre-online community	
		y	n
Coin *asexual*	y	40	17
	n	40	109

Table 6.1: Responses to "Did you ever use the term *asexual* to describe yourself before encountering online asexual communities?" and "As a term applying to people, did you think of the term *asexual* on your own?" for respondants with a yes/no response for each (n=206)

To test the hypothesis that seeing asexuality as a lack of sexuality is more common among asexuals who identified as asexual after thinking of the term on their own, I calculated the proportion of *yes* responses to the lack-of-sexuality item for those who answered *yes* to both of the questions reported in Table 6.1 and for those who answered *no* to at least one of those questions. Table 6.2 shows the responses to the lack-of-sexuality item for both groups when *other* responses that could not be back-coded are excluded.

		Is their "asexuality" a lack of "sexuality"?	
		y	n
Independently coin and	y	21	19
identify as *asexual*	n	53	93

Table 6.2: Participants' responses to the lack-of-sexuality item, divided according to a) those who said that they had both independently coined/adapted the term '*asexual* and identified as *asexual* before finding an online asexual community, and b) those who indicated *no* to at least one of those items.

The proportions of *yes* responses for those who independently thought of the term *asexual* and who identified as asexual prior to finding an online community (52.5%) was higher than for those who did not (36.3%). To test for statistical significance, I performed a Fisher's exact test (p = 0.1, OR = 1.89). The raw numbers indicate a difference of about 15 percentage points in the predicted direction, although the evidence against the null hypothesis is fairly weak due to the small (sub)sample size.

Participants were asked if they have ever identified as *nonsexual*. If they answered yes (n=95), they were given the same set of questions as for asexual along with a question asking if they currently use the term nonsexual to describe themselves. Table 6.3 shows the breakdown of whether they used the term *nonsexual* to describe themselves before finding

an online asexual community at the time of the survey (excluding *other* responses that could not be backcoded).

		Nonsexual before finding online community	
		y	n
Nonsexual now	y	15	30
	n	7	33

Table 6.3: Paired responses to the items "Did you ever use the term *nonsexual* to describe yourself before encountering online asexual communities?" and "Do you presently use the term *nonsexual* to describe yourself?"

The AVEN Shift hypothesis predicts that there should be a trend towards decreased identification as nonsexual after finding one or more online asexual communities. To test for this, I ran a McNemar test[9] for the *before* and *now* conditions ($p = 0.00019$), which strongly suggests that there was a difference in the expected direction (provided self report is accurate).

Another prediction of the AVEN shift hypothesis is that, through encountering online asexual communities, reading/watching relevant content (and possibly creating some), people's understanding of *asexual* should move away from "not sexual," move away from "not sexually active" and move towards "on analogy with hetero/homo/bisexual." I also expected that "not sexual" would be more strongly associated with *nonsexual*, and that "on analogy with hetero/homo/bisexual" would be more strongly associated with *asexual*. Because participants were told to check all that applied, I am treating these as separate (paired) yes/no questions. The proportions of *yes* responses for each meaning (for when people first considered themselves asexual, and when taking the survey) are shown in Figure 6.2. Responses for *nonsexual* are shown in Figure 6.3.

McNemar tests were performed on each pair (excluding "other"). For *asexual*, there was significant decrease in interpreting it as "not sexual" (34 yes/no vs. 14 no/yes. $p = 0.0055$), a decrease in interpreting it as "not sexually active" (28 yes/no vs. 4 no/yes. $p = 1.93e-05$), and an increase in interpreting it in terms of analogy (7 yes/no vs. 29 no/yes. $p = 0.0003$).

---

[9]McNemar tests are for paired binary responses on comparable binary tests (e.g. if a person has a condition or not before vs. after some treatment). For each individual, there are four possible response patterns: yes/yes, yes/no, no/yes, no/no. The null hypothesis is that p(yes/no) = p(no/yes). Because the number of relevant responses was generally low, I used a two-tailed binomial test.
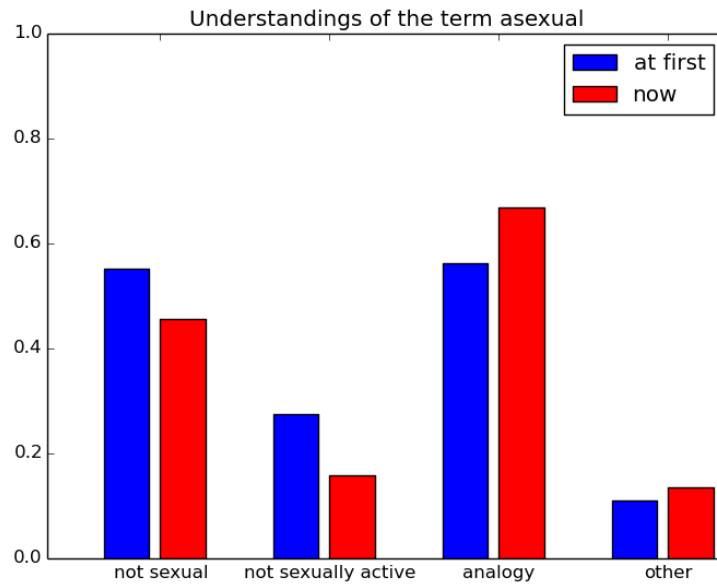
Figure 6.2: Proportion of *yes* responses for the questions "In what sense(s) did you understand *asexual* when first considering yourself asexual?" and for "In what sense(s) do you presently understand *asexual*?" Participants were told to check all that applied, so checking a box is interpreted as a *yes*, and not checking it as a *no*. In all of the 208 completed surveys, at least one box was checked for each question.

For *nonsexual*, there was a significant decrease in interpreting it as "not sexually active' (12 yes/no vs. 1 no/yes. $p = 0.0034$).

Participants were also asked about the terms *celibate* and *abstinent*. For each, they were presented a page asking the question: "Have you ever considered yourself *celibate* [abstinent]? (This is not asking about whether or not you are or were sexually active, but whether you have used the term celibate [abstinent] to describe yourself.)" For each term, if the participants answered *yes*, they were taken to a page asking about the term, such as whether they considered themselves celibate and/or abstinent before finding an online asexual community, after finding an online asexual community, and at present. A total of 75 participants indicated having ever considered themselves celibate, and 50 abstinent, and there was considerable overlap between these groups (n = 32; OR = 4.67, when restricting analysis to the 206 respondents who answered both questions). For *celibate*, 64 respondents gave usable answers for all three time periods, as did 45 of the respondents for *abstinent*.[10]

---

[10]This includes those that were back-coded, which were cases of "yes, but..." The other *other* responses tended to indicate

119
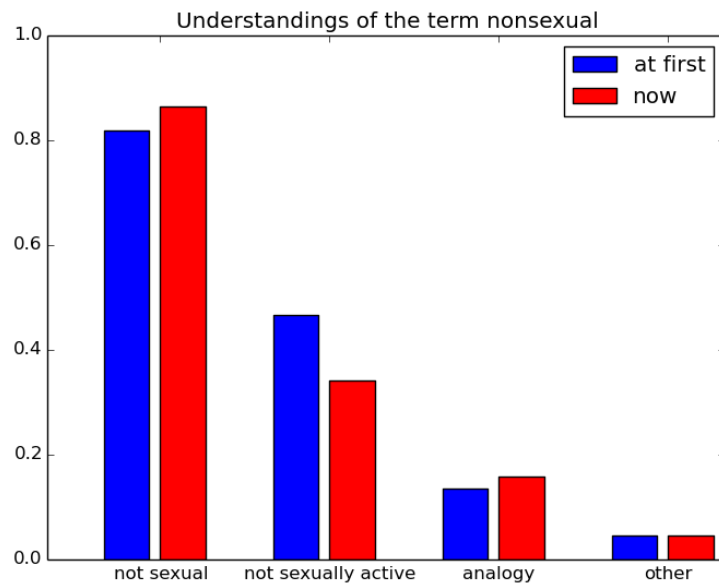
Figure 6.3: Proportion of *yes* responses for the questions "In what sense(s) did you understand *nonsexual* when first considering yourself nonsexual?" and for "In what sense(s) do you presently understand *nonsexual*?" Participants were told to check all that applied, so checking a box is interpreted as a *yes*, and not checking it as a *no*; 7 individuals did not check any boxes for at least one of the questions and were excluded for this analysis, leaving n = 88.
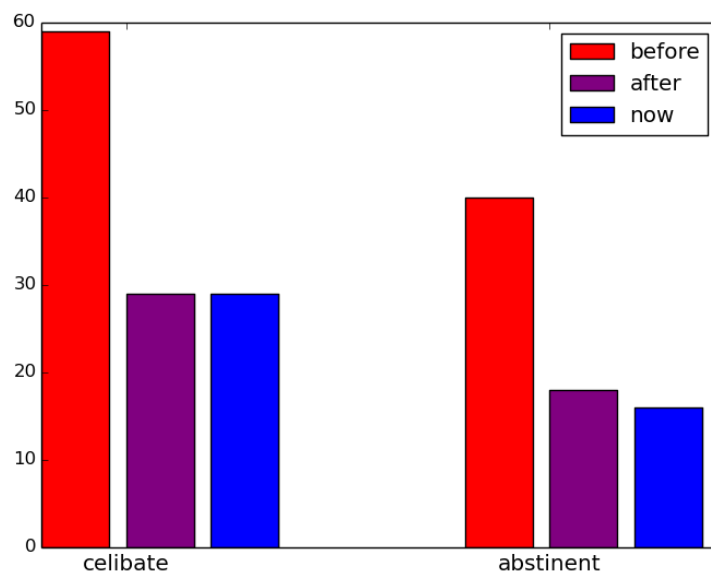
Figure 6.4: Number of affirmative responses to the questions "Did you ever use the term *celibate* to describe [*abstinent*] yourself before encountering online asexual communities?", "After finding an online asexual community, did you consider yourself celibate [abstinent]?" and "Do you still consider yourself *celibate*? [*abstinent*]" Data from respondants with at least one non-backcodable other:specify response are excluded (n=9 for *celibate*; n=5 for *abstinent*)

The number of *yes* responses for each question are shown in Figure 6.4 To test for an overall difference from the first to last time period, McNemar tests were performed, which revealed a significant decrease in regarding oneself as celibate (33 yes/no vs. 3 no/yes. $p = 2.27e - 7$) and also a significant decrease for abstinent (25 yes/no vs. 1 no/yes. $p = 8.05e - 7$).[11]

Figure 6.4 suggests that this change was overwhelmingly from *before* to *after*, with little change from *after* to *now*. Closer inspection of individual response patterns confirms this. From *before* to *after*, there were 33 yes/no response patterns for *celibate* and 23 for *abstinent* versus 3 and 1 no/yes response patterns, respectively. From *after* to *now*, for *celibate* there were 2 yes/no and 2 no/yes patterns, while for *abstinent* there were only two switches, both of which were yes/no.

things like "maybe", "sort of", or "yes and no". In some cases, the other:specify option was used as an opinion box.

[11]All response patterns were monotonic (i.e. There were no cases of yes/no/yes or of no/yes/no)
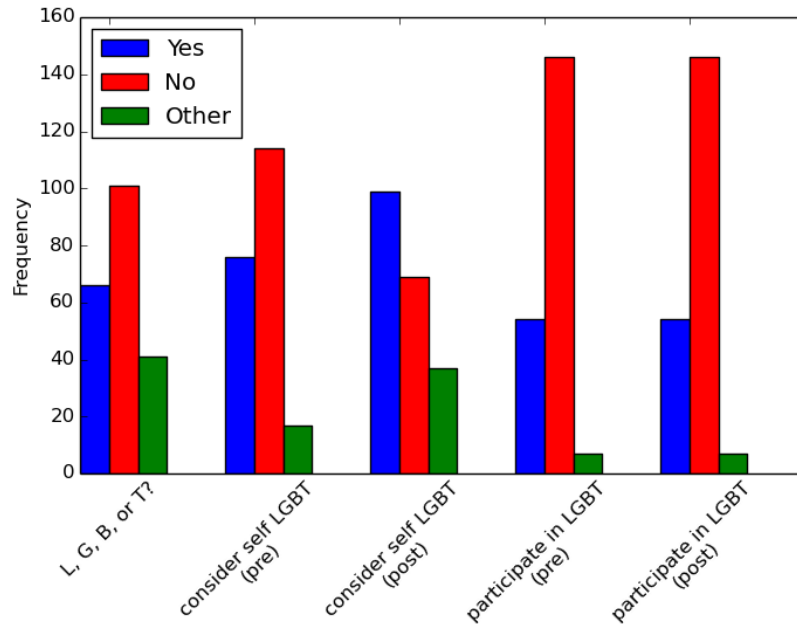
Figure 6.5: Raw counts for responses to the following questions:

"In addition to being asexual, are you gay, lesbian, bi, or transgender?"

"Prior to (Subsequent to) finding an online asexual community, did you consider yourself part of LGBT/queer?"

"Prior to (Subsequent to) finding an online asexual community, have you participated in one or more offline LGBT/queer communities?"

6.3.2 Relationship with LGBT

Participants were asked several questions concerning asexuality and LGBT identities and organizations. Respondents were asked whether, in addition to being asexual, they were gay, lesbian, bi, or transgender; whether they considered themselves part of LGBT/queer prior to finding an online asexual community and subsequent to finding one; and whether they participated in one or more offline LGBT/queer communities prior to finding online asexual community and subsequent to finding one. All had questions had *yes*, *no*, and *other* as possible responses. It was felt that few other responses could be back-coded, and so no back-coding was done. Raw scores are displayed in Figure 6.5

Because questions were on the same page, *pre* vs. *post* differences can be taken to indicate that the person believes there was a change, which I will interpret as an actual change. In

terms of offline involvement, the change from *pre* to *post* was fairly small (a decrease of 4.8 percentage points for *no*), but the change was more substantial for whether participants considered themselves part of LGBT (11 percentage point increase for yes and a 9.6 point increase for other). If this change was caused by finding an online asexual community, it suggests that finding such a community has a relatively small impact on people's off-line interactions with LGBT groups, but has a more substantial impact on people's perception of the matter.

For two questions, there was an especially high rate of *other*. The first of these was the question "In addition to being asexual, are you gay, lesbian, bi, or transgender?" The other responses usually involved a non-binary gender identity (e.g. *genderqueer*, *androgynous*, *neutrois*, etc.), romantic orientation involving at least some same-gender attraction (e.g. *biromantic*, *homoromantic*, *panromantic*), or uncertainty. Also having a high rate of *other* was the question, "Subsequent to finding an online asexual community, have you considered yourself part of LGBT/queer?" Some other responses were because of general uncertainty, and a few people chose other to indicate they are allies. However, most of the others reflected the uncertain relationship between asexuality and LGBT, suggesting that finding and participating in online asexual communities makes people less likely to not consider themselves part of LGBT, with some of these people feeling that they are part of LGBT and others feeling ambivalent about it.

6.3.3   Final Comments

At the end of the survey, participants were thanked for their participation and asked, "Before finishing this survey, is there anything you would like to say?" The main purpose of this was to allow feedback in case participants felt that survey questions were confusing, or if they they felt forced to misrepresent themselves because none of the options was good. Only one person criticized the questions, and most feedback was positive or neutral. The relative lack of negative feedback adds confidence to the validity of the data.

6.4 Discussion

There are numerous limitations of this survey. First, the data was mostly from people at the youngest end of the eligible demographic and the sample sizes for some of the older demographic were too small to do meaningful analysis for comparing age groups. This issue was compounded by using five year age groups rather than one year groups. Second, the data were based on retrospective reporting, and how accurate people's memories are is likely variable. Third, the number of years between the founding of currently existing online asexual communities and the collection of this data makes them far from ideal for understanding how self-identified asexuals understood asexuality prior to finding online asexual communities. Fourth, this is a convenience sample, and the median for finding an online asexual community was 2009 and for joining one was 2010, suggesting that the data are skewed towards more recent participation.

Despite these limitations, a number of findings have emerged. First, the results generally confirm the five hypotheses stemming from the theory that there has been an *AVEN shift*, although the changes are far from monolithic and larger in some cases than in others. Also, one of the hypotheses (people's understanding of asexuality should move in the direction of sexual orientation and seeing themselves aligned with LGBT) should be qualified by noting that a large part of this shift involves people moving from not seeing themselves as part of LGBT to feeling ambivalent about it. While survey results cannot prove either that these changes were caused by finding online asexual communities, or as a direct (or indirect) result of the influence of AVEN, finding an online asexual community seems to be the most plausible explanation, especially in conjunction with the historical evidence discussed in Chapter 4.

Chapter 7

General Discussion and Conclusion

The purpose of this dissertation is to examine the development of, and changes in, online English language asexual discourse from the second half of the 1990s until late 2013. The primary focus is on lexical items. This study provides a contemporary case study for a microcosm of language evolution and develops novel methodological approaches for sociolinguistics and corpus linguistics for studying online language change in specific communities.

After reviewing previous works on language and sexuality, lexical change, and language evolution (Chapter 2), I surveyed both historical and relatively recent academic discourse about asexuality (Chapter 3) as part of an examination of options available for conceptualizing asexuality. In Chapter 4, I further explored available conceptual options through searching through newspaper archives. I then give a historical account of the development of online asexual communities, and I argue that there have been at least two major conceptual shifts in the conceptualization of asexuality in the time period under consideration: the *AVEN shift* and the *rise of intermediate categories*.

The specific research questions addressed in this dissertation are:

1. Is there anything that distinguishes asexual discourse from other kinds of discourse?

2. If yes, what words are most characteristic of asexual discourse, and what can these tell us about asexual discourse?

3. Are expressions of undesire and/or nondesire important parts of online asexual discourse?

4. Does quantitative research support the *AVEN shift* and *rise of intermediate categories* hypotheses?

7.1 Research and approaches

This dissertation has adopted three major research approaches—archival research (Chapters 3 and 4), corpus research (Chapter 5), and survey research (Chapter 6). Using the Way Back Machine, databases of newspaper archives, and academic databases and references in papers, historical conceptualizations of asexuality can be seen well before the emergence of asexual communities online. Corpus linguistic methods can provide large scale text data for the analysis of lexical use, lexical development, and keywords analysis. A major limitation of using corpus data for many online communities (including the ones studied here) is that the data is demographically lean, i.e. the corpora cannot provide users' demographic information for the study. Furthermore, because the *AVEN shift* hypothesis largely involves a) changes from before vs. after finding online asexual communities, and b) changes in asexual discourse during the times when texual data is the most sparse, the asexual corpora are not well-suited for testing this hypothesis. Using an online survey approach, I was able to obtain data better suited to testing the *AVEN shift* hypothesis. This dissertation combined these three research approaches to build a integrated research frame to serve the goals of this dissertation.

7.2 Findings

I constructed four corpora (Chapter 5). By using a multinomial Naive Bayes classifier, I found moderate distinguishability between the AVEN-main and non-asexual corpra at the level of individual posts when only considering individual words. In order to rule out the possibility that, perhaps, the classifier was distinguishing AVEN vs. non-AVEN discourse (rather than asexual vs. non-asexual discourse), I designed two algorithms to automatically detect words likely referring to users. These, along with some manually-identified names were ignored by the classifier. In addition, I used the same classifier on the AVEN-other corpora and the asexual-other corpora. Results for the Asexual-other corpus are similar to results for the AVEN-main, while results for the AVEN-other corpus are not. This suggests

that the classier is identifying asexual discourse vs. other discourse.

Moreover, I used the AVEN-core corpus to generate a list of key-words that well-characterize asexual discourse. A well-known problem in corpus linguistics is how to exclude extremely "clumped" words from these lists. Existing approaches in corpus linguistics do not work well for message board data because the number of words per part is not a balanced corpus: Posts per user follows a power law, and words-per user is strongly correlated with posts-per-user. I propose using a different metric of dispersion, which is more appropriate for data from online message boards (Appendix B).

The keyword analysis found several common themes, including various forms of or words derived from *asexual*, and two roots were especially prominent *sex* (typically involving patterns of attraction or attitudes about sex) and *romance/romantic*. Some of the other most common themes involved a) sex and/or sexuality, b) attraction, emotions, and desire/undesire/nondesire, c) relationships.

Included among these were several intermediate-category terms (e.g. *gray-a*, *demisexual*, and *hyposexual*), suggesting the importance of these categories. The corpus data is date-stamped, and this allowed for examination of various areas of asexuality-related terminology, including intermediate-category terms. Results showed that, since the earliest period of the data, there has been use of intermediate-category terms. Initially, the most popular was *hyposexual*. Around 2007–2008, the terms *gray-A* and *demisexual* became much more common. The use of *hyposexual* declined considerably, while *gray-A* and *demisexual* each acquired far more currency than *hyposexual* ever did. This provides strong confirmation of the *rise of intermediate-categories* hypothesis.

To test the *AVEN shift* hypothesis, in early 2012 I conducted an online survey for asexual-identified individuals who were at least 25 years of age and had lived most of their lives since age 12 in English language contexts. The *AVEN shift* hypothesis makes several testable hypotheses about people's self-identification and their interpretation of various terms before vs. after finding an online asexual community:

1. There should be a decrease in the number of people identifying as *nonsexual.*

2. There should be a decrease in people considering themselves *abstinent* or *celibate.*

3. People's understanding of asexuality should move in the direction of "not experiencing sexual attraction."

4. People's understandings of asexuality should move away from "not sexual" or "lacking sexuality."

5. People's understanding of asexuality should move in the direction of "sexual orientation" and seeing themselves aligned with LGBT.

Results from the survey provided evidence supporting all five predictions, although the changes were by no means monolithic. This provides quantitative support for the *AVEN shift* hypothesis.

7.3   Applications and further research

The study of the evolution of online asexual discourse in this dissertation provides a contemporary case study in the ongoing development of a kind of discourse and lexical change in which a huge part of the language involved in doing so is machine readable and in the public domain. It demonstrates the potential value of using data from online message boards for investigating recent lexical change and discourse development, and it makes novel contributions to corpus linguistics methodology, which are likely to be important for anyone endeavoring to do similar research on other online communities.

First, one challenge researchers will face is the automatic detection of names referring to users. While this was not a major focus of this dissertation, I provide two approaches for doing this. Future research would benefit from detailed analysis of these or other means of identifying names for users. Possible improvements for these algorithms include a) applying more sophisticated approaches to finding sentence boundaries, and b) using a morphological parser to provide possible diminutive nicknames for users.

Second, generating a list of keywords should use some kind of dispersion metric to exclude extremely clumped words that are not representative of the discourse as a whole. Nearly all

of the main parts-based metrics rely on the assumption that the data is a balanced corpus (or at least approximately balanced). Actual data in the online message boards investigated here are not remotely balanced—a large proportion of users post only one or a few times, while some users are extremely prolific. Altmann et al. (2011) use a dispersion metric (for message board textual data), which I have used in generating my list of keywords, and this approach would work well for other researchers using message-board data.

The increasing, world-wide use of a host of forms of social media provides researchers with access to historically unprecedented amounts of textual data, and this data is especially well-suited for studying lexical and discourse evolution. In this dissertation, I hope I have provided some useful methodological ideas for what kinds of research are possible in this area, and provided answers to some of the challenges other researchers are likely to face.

Chapter 8

References

Agarwal, A. (2013). Download the old archives of your website from google reader before it's too late. Retrieved 4/23/2013 from http://www.labnol.org/internet/download-website-archives/27981/.

Aicken, C. R., Mercer, C. H., and Cassell, J. A. (2013). Who reports absence of sexual attraction in britain? evidence from national probability surveys. *Psychology & Sexuality*, 4(2):121–135.

Altmann, E. G., Pierrehumbert, J. B., and Motter, A. E. (2009). Beyond word frequency: Bursts, lulls, and scaling in the temporal distributions of words. *PLoS One*, 4(11):e7678.

Altmann, E. G., Pierrehumbert, J. B., and Motter, A. E. (2011). Niche as a determinant of word fate in online groups. *PloS one*, 6(5):e19009.

Alvarez, W. C. (1959a). Impotence in hubby wife's worry. *The Miami News. June 15, 1959.*, page 18c.

Alvarez, W. C. (1959b). Men who would be women victims of cruel hoax.

American Psychiatric Association (1952). *Diagnostic and statistical manual of psychiatric disorders.* American Psychiatric Pub.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders (2nd edition).* American Psychiatric Pub.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders (3rd edition).* American Psychiatric Pub.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders (3rd edition-revised).* American Psychiatric Pub.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th edition).* American Psychiatric Pub.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders (4th edition-text revision).* American Psychiatric Pub.

Baker, P. (2002). *Polari-the lost language of gay men.* Routledge.

Baker, P. (2004). Querying keywords questions of difference, frequency, and sense in keywords analysis. *Journal of English Linguistics*, 32(4):346–359.

Baker, P. (2005). *Public discourses of gay men.* Routledge.

Baker, P. (2008). *Sexed texts: language, gender and sexuality.* Equinox.

Barrett, R. (2002). Is queer theory important for sociolinguistic theory. In *Language and sexuality: Contesting meaning in theory and practice*, pages 25–43.

Bell, A. (1978). Asexual chic: Everybody's not doing it. *The Village Voice. January 23, 1978*, pages 1, 20–21.

Bentley, R. A. (2008). Random drift versus selection in academic vocabulary: An evolutionary analysis of published keywords. *PloS one*, 3(8):e3057.

Bentley, R. A., Garnett, P., O'Brien, M. J., and Brock, W. A. (2012). Word diffusion and climate science. *PloS one*, 7(11):e47966.

Berkey, B. R., Perelman-Hall, T., and Kurdek, L. A. (1990). The multidimensional scale of sexuality. *Journal of Homosexuality*, 19(4):67–88.

Billington, M. (1971). Ballad of the sad café (review). *The Times. March 10, 1971*, page 11.

Bird, S., Klein, E., and Loper, E. (2009). *Natural language processing with Python.* " O'Reilly Media, Inc.".

Bishop, C. (2013). A mystery wrapped in an enigma–asexuality: a virtual discussion. *Psychology & Sexuality*, 4(2):195–206.

Blanchard, R. (1985). Typology of male-to-female transsexualism. *Archives of Sexual Behavior*, 14(3):247–261.

Blanchard, R. (1989a). The classification and labeling of nonhomosexual gender dysphorias. *Archives of sexual behavior*, 18(4):315–334.

Blanchard, R. (1989b). The concept of autogynephilia and the typology of male gender dysphoria. *The Journal of nervous and mental disease*, 177(10):616–623.

Blommaert, J. (2005). *Discourse: A critical introduction.* Cambridge University Press.

Bogaert, A. F. (2003). Number of older brothers and social orientation: New tests and the attraction/behavior distinction in two national probability samples. *Journal of Personality and Social Psychology*, 84(3):644.

Bogaert, A. F. (2004). Asexuality: Prevalence and associated factors in a national probability sample. *Journal of Sex Research*, 41(3):279–287.

Bogaert, A. F. (2006). Toward a conceptual understanding of asexuality. *Review of General Psychology*, 10(3):241.

Bogaert, A. F. (2008). Asexuality: Dysfunction or variation. In *Psychological sexual dysfunctions*, pages 9–13. Nova Science. New York, NY.

Bogaert, A. F. (2012a). Asexuality and autochorissexualism (identity-less sexuality). *Archives of sexual behavior*, 41(6):1513–1514.

Bogaert, A. F. (2012b). *Understanding asexuality.* Rowman & Littlefield Publishers.

Bogaert, A. F. (2013). The demography of asexuality. In *International Handbook on the Demography of Sexuality*, pages 275–288. Springer.

Bogaert, A. F. and Friesen, C. (2002). Sexual orientation and height, weight, and age of puberty: New tests from a british national probability sample. *Biological Psychology*, 59(2):135–145.

Bolton, R. (1995). Sex talk: Bodies and behaviors in gay erotica. In *Beyond the Lavender Lexicon. Authenticity, Imagination, and Appropriation in Lesbian and Gay Languages. Gordon and Breach Publishers*, pages 173–206.

Boyd, M. (1971). A priest says, 'it doesn't have a soul.'. *New York Times. October 24, 1971*, page D1.

Brotto, L. A., Knudson, G., Inskip, J., Rhodes, K., and Erskine, Y. (2010). Asexuality: A mixed-methods approach. *Archives of Sexual Behavior*, 39(3):599–618.

Brotto, L. A. and Yule, M. A. (2009"). Reply to hinderliter (2009). *Archives of Sexual Behavior*, 38(5):622–623.

Brotto, L. A. and Yule, M. A. (2011). Physiological and subjective sexual arousal in self-identified asexual women. *Archives of sexual behavior*, 40(4):699–712.

Bucholtz, M. and Hall, K. (2004). Theorizing identity in language and sexuality research. *Language in Society*, 33(04):469–515.

Bullough, V. L. (1994). *Science in the bedroom: A history of sex research*. Basic Books.

Cameron, D. (2005). Language, gender, and sexuality: Current issues and new directions. *Applied Linguistics*, 26(4):482–502.

Cameron, D. and Kulick, D. (2003). *Language and sexuality*. Cambridge University Press.

Cameron, D. and Kulick, D. (2005). Identity crisis? *Language & Communication*, 25(2):107–125.

Cameron, D. and Kulick, D. (2006). *The language and sexuality reader*. Taylor & Francis.

Campbell-Kibler, K. (2002). *Language and sexuality: contesting meaning in theory and practice*. Center for the Study of Language and Information.

Carrigan, M. (2011). There's more to life than sex? difference and commonality within the asexual community. *Sexualities*, 14(4):462–478.

Carrigan, M. A. (2012). How do you know you don't like it if you haven't tried it? In *Sexual Minority Research in the New Millennium*, pages 3–20.

Cerankowski, K. J. and Milks, M. (2010). New orientations: Asexuality and its implications for theory and practice. *Feminist Studies*, pages 650–664.

Chasin, C. D. (2011). Theoretical issues in the study of asexuality. *Archives of sexual behavior*, 40(4):713–723.

Church, K. W. and Gale, W. A. (1995). Poisson mixtures. *Natural Language Engineering*, 1(02):163–190.

CNN (2004). Study: One in 100 adults asexual. *CNN. Retrieved 8/21/2012 from http://www.cnn.com/2004/TECH/science/10/14/asexual.study/index.html*.

Conner, K. (1996). *Covariation of sexual orientation and sexual desire*. PhD thesis, University of Florida.

Conner, R. P. (1997). Les molles et les chausses: Mapping the isle of hermaphrodites in premodern france. In *Queerly phrased: Language, gender, and sexuality*, volume 127, page 46.

Davies, N. (1983). Nilsen confessed to killing and cutting up 16 men. *The Guardian. October 25. 1983*, pages 1,3.

Eckert, P. (2002). Demystifying sexuality and desire. *Language and sexuality: Contesting meaning in theory and practice*, pages 99–110.

English, R. (1998). *Leather Spinsters and the Degrees of Asexuality*. Author.

Erens, B., McManus, S., Field, J., Korovessis, C., Johnson, A., Fenton, K., and Wellings, K. (2001). *National survey of sexual attitudes and lifestyles II: Technical report*. National Centre for Social Research London.

Flaherty, J. (1971). Asexuals have problems too. *The Village Voice. February 25, 1971*, page 21.

Foucault, M. (1978/1995). *The history of sexuality, volume 1: An introduction (R. Hurley, Trans.)*. New York: Pantheon.

Francis, A. M. (2013). The wages of sin: How the discovery of penicillin reshaped modern sexuality. *Archives of sexual behavior*, 42(1):5–13.

Gazzola, S. B. and Morrison, M. A. (2011). Asexuality: An emergent sexual orientation. *Sexual minority research in the new millennium. Hauppauge, NY: Nova Science*.

Gee, J. (1991). Socio-cultural approaches to literacy (literacies). *Annual review of applied linguistics*, 12:31–48.

Ginoza, M. K., Miller, T., and AVEN Survey Team (2014). The 2014 aven community census: Preliminary findings.

Gochros, H. L. and Gochros, J. S. (1977). *The sexually oppressed*. Association Press.

Good, I. J. (1950). Probability and the weighing of evidence.

Goodman, S. N. (1999). Toward evidence-based medical statistics. 2: The bayes factor. *Annals of internal medicine*, 130(12):1005–1013.

Gressgård, R. (2013). Asexuality: from pathology to identity and beyond. *Psychology & Sexuality*, 4(2):179–192.

Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. *International journal of corpus linguistics*, 13(4):403–437.

Guardian, T. (1969). Miscellany. *The Guardian. July 29, 1969*, page 9.

Haeberle, E. J. (1998). Bisexuality: History and dimensions of a modern scientific problem. In *Bisexualities - The Ideology and Practice of Sexual Contact with both Men and Women*, pages 14–51.

Haefner, C. (2011). *Asexual scripts: A grounded theory inquiry into the intrapsychic scripts asexuals use to negotiate romantic relationships*. PhD thesis, Institute of Transpersonal Psychology.

Hallucigenia (2008). Letter from the editor. *AVENues*, page 2. Retrieved 4/27/2014 from http://www.asexuality.org/avenues/2008_08_16.pdf.

Herdt, G. (1991a). Representations of homosexuality: An essay on cultural ontology and historical comparison part i. *Journal of the History of Sexuality*, pages 481–504.

Herdt, G. (1991b). Representations of homosexuality: an essay on cultural ontology and historical comparison part ii. *Journal of the History of Sexuality*, pages 603–632.

Hinderliter, A. C. (2009a). Asexuality: The history of a definition.

Hinderliter, A. C. (2009b). Methodological issues for studying asexuality. *Archives of sexual behavior*, 38(5):619–621.

Hinderliter, A. C. (2013). How is asexuality different from hypoactive sexual desire disorder? *Psychology & Sexuality*, 4(2):167–178.

Hirschfeld, M. (1896). *Sappho und Sokrates: oder, Wie erklärt sich die Liebe der Männer und Frauen zu Personen des eigenen Geschlechts? (Sappho and Socrates: How is the love of men and women to persons of their own sex to be explained?)*. M. Spohr.

Hock, H. H. (1991). *Principles of historical linguistics*. Walter de Gruyter.

Höglund, J., Jern, P., Sandnabba, N. K., and Santtila, P. (2014). Finnish women and men who self-report no sexual attraction in the past 12 months: Prevalence, relationship status, and sexual behavior history. *Archives of sexual behavior*.

Hoppe, A. (1979). Asexual militancy. *Spokane Daily Chronicle. November 10, 1979*, page 5.

Horn, L. (1989). *A natural history of negation*. IL: University of Chicago Press, Chicago.

Horn, L. (2011). Google ending newspaper archiving project. Retrieved 4/29/2013 from http://www.pcmag.com/article2/0,2817,2385664,00.asp.

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing In Science & Engineering*, 9(3):90–95.

Iorio, J. B. (2011). Explaining orthographic variation in a virtual community: Linguistic, social, and contextual factors.

Irvine, J. M. (1990/2005). *Disorders of desire: sexuality and gender in modern American sexology.* Temple University Press.

Jackman, S. (2009). *Bayesian analysis for the social sciences*, volume 846. John Wiley & Sons.

Jay, D. (2003). A look at online collective identity formation.

Jay, D. (2006). History lesson. from http://asexualunderground.blogspot.com/2006/08/5-history-lesson.html. Retrieved 10/1/2007.

Jespersen, O. (1922). Language: its nature, development and origin.

Johnson, M. T. (1977). Asexual and autoerotic women: Two invisible groups. *The sexually oppressed*, pages 96–109.

Kaplan, H. S. (1977). Hypoactive sexual desire. *Journal of Sex & Marital Therapy*, 3(1):3–9.

Keller, R. (1994). *On language change: The invisible hand in language (3rd edition).* Routledge.

Kilgarriff, A. (2009). Simple maths for keywords. In *Proc. Corpus Linguistics*.

Kilgarriff, A. (2012). Getting to know your corpus. In *Text, Speech and Dialogue*, pages 3–15. Springer.

Kim, E. (2010). How much sex is healthy? the pleasures of asexuality. In Kirkland, A. and Metzl, J., editors, *Against health: How health became the new morality*, pages 157–169. New York University Press New York.

Kim, E. (2011). Asexuality in disability narratives. *Sexualities*, 14(4):479–493.

Kinsey, A. C., Pomeroy, W. B., and E, M. C. (1948). Sexual behavior in the human male.

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., and Gebhard, P. H. (1953). *Sexual behavior in the human female.* W. B. Saunders Company.

Kleinfeld, M. S. and Warner, N. (1997). Lexical variation in the deaf community relating to gay, lesbian, and bisexual signs. In *Queerly phrased: Language, gender, and sexuality*, pages 58–84.

Krafft-Ebing, R. (1903/1998). *Psychopathia sexualis: A medico-forensic study.* Putnam.

Kulick, D. (1999). Language and gender/sexuality. In *Sixth Language and Culture Online Symposium. http://www. language-culture. org/colloquia/symposia/kulick-don.*

Kulick, D. (2000). Gay and lesbian language. *Annual Review of Anthropology*, pages 243–285.

Kulick, D. (2002). Queer linguistics. *Language and sexuality: Contesting meaning in theory and practice*, pages 65–68.

Kulick, D. (2008). 5 language and desire. *The handbook of language and gender*, 25:119.

Kulick, D. (2009). *Travesti: Sex, gender, and culture among Brazilian transgendered prostitutes.* University of Chicago Press.

Lakoff, R. (1973). Language and woman's place. *Language in society*, 2(01):45–79.

Landers, A. (1979). Asexual adult reveals feelings. *Toledo Blade. April 16, 1979*, page 21.

Leap, W. (1995). *Beyond the lavender lexicon: authenticity, imagination, and appropriation in lesbian and gay languages.* Taylor & Francis US.

Leap, W. L. and Motschenbacher, H. (2012). Launching a new phase in language and sexuality studies. *Journal of Language and Sexuality*, 1(1):1–14.

Lee, J. J. (2010). Mechanize. http://wwwsearch.sourceforge.net/mechanize/.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., and Wolff, S. (2009). A brief history of the internet. *ACM SIGCOMM Computer Communication Review*, 39(5):22–31.

Lever, J., Kanouse, D. E., Rogers, W. H., Carson, S., and Hertz, R. (1992). Behavior patterns and sexual identity of bisexual males. *Journal of Sex Research*, 29(2):141–167.

Lewis, D. D. (1998). Naive (bayes) at forty: The independence assumption in information retrieval. In *Machine learning: ECML-98*, pages 4–15. Springer.

Lief, H. I. (1977). Inhibited sexual desire. *Medical Aspects of Human Sexuality*, 7(1).

Livia, A. (2002). Camionneuses s' abstenir: Lesbian community creation through the personals. In *Language and sexuality: Contesting meaning in theory and practice*, pages 191–206. California Center for the Study of Language and Information Stanford.

Livia, A. and Hall, K. (1997a). It's a girl: Bringing performativity back to linguistics. In *Queerly Phrased: Language, gender and sexuality*, pages 3–18. Oxford University Press.

Livia, A. and Hall, K. (1997b). *Queerly phrased: Language, gender, and sexuality.* Oxford University Press.

Lucassen, M. F., Merry, S. N., Robinson, E. M., Denny, S., Clark, T., Ameratunga, S., Crengle, S., and Rossen, F. V. (2011). Sexual attraction, depression, self-harm, suicidality and help-seeking behaviour in new zealand secondary school students. *Australian and New Zealand journal of psychiatry*, 45(5):376–383.

Manalansan, M. F. (1995). Performing'the filipino gay experiences in america: linguistic strategies in a transnational context. In *Beyond the Lavender Lexicon. Newark, NJ: Gordon and Breach*, pages 249–66.

Masters, W. H. and Johnson, V. E. (1970). *Human sexual inadequacy*, volume 225. Boston Little, Brown, and Company.

McClave, C. H. (2013). Asexuality as a spectrum: A national probability sample comparison to the sexual community in the uk. Master's thesis, Columbia University.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., et al. (2011). Quantitative analysis of culture using millions of digitized books. *science*, 331(6014):176–182.

Miller, T. (2012). Analysis of the 2011 asexual awareness week community census.

Moorhouse, G. (1969). Lawrence of many truths. *The Guardian. August 11, 1969*, page 10.

Morrish, L. and Sauntson, H. (2007). *New perspectives on language and sexual identity*. Palgrave Macmillan Basingstoke.

Murphy, J., Hashim, N. H., and O'Connor, P. (2007). Take me back: validating the wayback machine. *Journal of Computer-Mediated Communication*, 13(1):60–75.

Murphy, M. (1997). The elusive bisexual: Social categorization and lexico-semantic change. In *Queerly phrased: Language, gender, and sexuality*, pages 35–57.

Nurius, P. S. (1983). Mental health implications of sexual orientation. *Journal of Sex Research*, 19(2):119–136.

Nuttbrock, L., Bockting, W., Mason, M., Hwahng, S., Rosenblum, A., Macri, M., and Becker, J. (2011). A further assessment of blanchard's typology of homosexual versus non-homosexual or autogynephilic gender dysphoria. *Archives of sexual behavior*, 40(2):247–257.

Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science & Engineering*, 9(3):10–20.

Pagel, M., Atkinson, Q. D., and Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout indo-european history. *Nature*, 449(7163):717–720.

Peng, F. and Schuurmans, D. (2003). Combining naive bayes and n-gram language models for text classification. In *Advances in Information Retrieval*, pages 335–350. Springer.

Peng, F., Schuurmans, D., Wang, S., and Keselj, V. (2003). Language independent authorship attribution using character level language models. In *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 1*, pages 267–274. Association for Computational Linguistics.

Pew Research Center (2015). Changing attitues on gay marriage. Retrieved 2016-01-31 from http://www.pewforum.org/2015/07/29/graphics-slideshow-changing-attitudes-on-gay-marriage/.

Pinker, S. (1994). *The language instinct: The new science of language and mind*, volume 7529. Penguin UK.

Porter, M. (1995). Talking at the margins: Kenyan discourses on homosexuality. In *Beyond the Lavender Lexicon: authenticity, imagination, and appropriation in lesbian and gay languages. New York NY: Gordon and Breach*.

Poston, D. L. and Baumle, A. K. (2010). Patterns of asexuality in the united states. *Demographic Research*, 23(18):509–530.

Prause, N. and Graham, C. A. (2007). Asexuality: Classification and characterization. *Archives of Sexual Behavior*, 36(3):341–356.

Przybylo, E. (2011). Crisis and safety: The asexual in sexusociety. *Sexualities*, 14(4):444–461.

Przybylo, E. (2013). Producing facts: Empirical asexuality and the scientific study of sex. *Feminism & Psychology*, 23(2):224–242.

Python Software Foundation (2010). Python 2.7. https://docs.python.org/2.7/.

Queen, R. (2007). Sociolinguistic horizons: Language and sexuality. *Language and Linguistics Compass*, 1(4):314–330.

Radford, T. (2004). No sex please, we're asexual. *The Guardian. October 13, 2004*.

Reiss, I. L. (2006). *An insider's view of sexual science since Kinsey*. Rowman & Littlefield.

Richardson, L. (2013). Beautiful soup. http://www.crummy.com/software/BeautifulSoup/.

Sapir, E. (1929/1992). *Male and female forms of speech in Yana*. California Indian Library Collections.

Scherrer, K. S. (2008). Coming to an asexual identity: Negotiating identity, negotiating desire. *Sexualities*, 11(5):621–641.

Scherrer, K. S. (2010a). Asexual relationships: What does asexuality have to do with polyamory. *Understanding non-monogamies*, pages 154–159.

Scherrer, K. S. (2010b). What asexuality contributes to the same-sex marriage discussion. *Journal of gay & lesbian social services*, 22(1-2):56–73.

Schiffrin, D. (1994). *Approaches to discourse*, volume 8. Blackwell Oxford.

Scott, M. (2014). Wordsmith tools manual, 6.0.

Shively, M. G. and De Cecco, J. P. (1977). Components of sexual identity. *Journal of homosexuality*, 3(1):41–48.

Smart, W. (1977). Is carter's image genuine? *The Desert News. March 11, 1977*.

Spence, J. T., Helmreich, R., and Stapp, J. (1975). Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of personality and social psychology*, 32(1):29.

Spitzer, R. L. and Wakefield, J. C. (1999). Dsm-iv diagnostic criterion for clinical significance: Does it help solve the false positives problem? *American Journal of Psychiatry*, 156(12):1856–1864.

Storms, M. D. (1979). Sexual orientation and self-perception. *Perception of Emotion in Self and Others*, pages 165–180.

Storms, M. D. (1980). Theories of sexual orientation. *Journal of Personality and Social Psychology*, 38(5):783.

Sundrud, J. L. (2011). Performing asexuality through narratives of sexual identity. Master's thesis, San José State University.

Terkourafi, M. (2011). The pragmatic variable: Toward a procedural interpretation. *Language in Society*, 40(03):343–372.

Titman, N. (2012). Asexuality bc (before cake).

Valentine, J. (1997). Pots and pans: Identification of queer japanese in terms of discrimination. In *Queerly phrased: Language, gender, and sexuality*, pages 95–114. New York: Oxford University Press.

Van Buren, A. (1981). Happy celibates want to stay that way. *The Bryan Times. May 4, 1981.*, page 7.

Van Houdenhove, E., Gijs, L., T'Sjoen, G., and Enzlin, P. (2014). Stories about asexuality: A qualitative study in asexual women. *Journal of sex & marital therapy*, (just-accepted).

Vrangalova, Z. and Savin-Williams, R. C. (2012). Mostly heterosexual and mostly gay/lesbian: Evidence for new sexual orientation identities. *Archives of sexual behavior*, 41(1):85–101.

Wadsworth, J., Field, J., Johnson, A., Bradshaw, S., and Wellings, K. (1993). Methodology of the national survey of sexual attitudes and lifestyles. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 407–421.

Wallis, L. (2012). What is it like to be asexual? *BBC News Magazine*, pages retrieved 3/14/2014 from https://web.archive.org/web/20120117175302/http://www.bbc.co.uk/news/magazine–16552173.

Watt, D. (1997). Read my lips: Clippyng and kyssyng in the early sixteenth century. In *Queerly phrased: Language, gender, and sexuality*, volume 167, page 77.

Weinreich, U., Labov, W., and Herzog, M. (1968). Empirical foundations for a theory of language change.

Westphal, S. (2004). Glad to be asexual. *New Scientist*, 17(7):40–43.

Wierzbicka, A. (1997). Understanding cultures through their key words: English, russian, polish, german, and japanese.

Wikimedia Commons (2012). Internet users per 100 inhabitants.

Wong, A. (2002). The semantic derogation of tongzhi: A synchronic perspective. *Language and sexuality*, pages 161–74.

Wong, A., Roberts, S. J., and Campbell-Kibler, K. (2002). Speaking of sex. *Language and Sexuality. Stanford: CSLI Publishing.*

Wong, A. and Zhang, Q. (2000). The linguistic construction of the tongzhi community. *Journal of Linguistic Anthropology*, 10(2):248–278.

Yule, M. A., Brotto, L. A., and Gorzalka, B. B. (2013). Mental health and interpersonal functioning in self-identified asexual men and women. *Psychology & Sexuality*, 4(2):136–151.

Yule, M. A., Brotto, L. A., and Gorzalka, B. B. (2014). Biological markers of asexuality: Handedness, birth order, and finger length ratios in self-identified asexual men and women. *Archives of sexual behavior*, 43(2):299–310.

Zwicky, A. (2012). Sexual clippings. Retrieved 3/19/2014 from http://arnoldzwicky.org/2012/12/09/sexual-clippings/.

Appendix A

Automatic detection of names referring to users

In order to automatically detect names referring to users, I developed two algorithms. Because automatic detection of names referring to users is not the focus of this dissertation, their accuracy was not rigorously tested. Obviously, the general goal is to maximize true positives and true negatives, while minimizing false positives and false negatives. More specifically, I was most interested in a) identifying names referring to users with moderate to high frequency, and b) avoid false positives that are, intuitively, important words for asexual discourse.[1]

However, looking through the list of words thus identified, both algorithms appear to work fairly well.

A.1 First algorithm

The first algorithm is based on the idea that a) given a user name, certain subsequences are especially likely to be options for how people address them, and b) if one of these strings has a strong tendency to be used near that user's posts, it likely refers to that user.

First, I ran the name through the same cleaner as words in the corpus (as this is how it would appear in the corpus). Usernames are unique prior to cleaning, but they are not necessarily unique after cleaning. Therefore, an uncleaned version is used as the author tag in the meta-data for each post. If there were no spaces, the only relevant string is the whole string. If there are spaces, then relevant subsequences are each word, as well as an initialism

---

[1]Simply removing a word if it is the name of a user would create serious problems. For example, after cleaning names the same way the text was cleaned, there is a user with the (cleaned) name *asexual* who has 4 posts. It would be a mistake to remove the word *asexual*

consisting of the first letter of each word. If a word belonged to a stopword list, it was excluded.[2]

For each relevant string, each occurrence in the forum was found. Then for each member for which some string was a potential name, the percent of uses of the string in the same thread within a certain window of posts as the possible user was calculated. I chose 60 posts on either side as the relevant window because this is two pages according to AVEN's default settings. (Note that this cut-off does not even apply in the large majority of threads, as there are fewer than 62 posts in 95.7% of the AVEN threads and 80.3% of the xkcd threads).

I then had to choose a threshold above which a string would be identified as a name. This was accomplished by looking through the results and seeing at around what point there began to be ordinary English words that were not part of the name of a well-known user. Based on this, I chose 30% as the cut-off.

This algorithm was effective for identifying a large number of usernames, but it has two major limitations. First, AVEN allows users to change their names, and so their name at the time of data collection may be different than their name when posting in a given thread. Further, some frequent posters are sometimes referred to by nicknames the above method would not identify, such as nicknames involving clippings, and possibly the addition of one or more morpheme (i.e. the diminutive suffix /i/, as in *mommy*). Therefore, I supplemented it with a second algorithm.

A.2   Second algorithm

The second algorithm is based on the following idea: Users are sometimes addressed via `@USERNAME`, but more commonly they are addressed with vocative-indicating punctuation. (e.g. "Fluffybunny47, how are you?" or "How are you, Fluffybunny47?")

Thus, in the uncleaned data, for each word in the corpus (other than those in a stopword list), the program calculated the percent of times that it was a) used sentence initially and

---

[2]The stopword list was based on http://www.textfilter.com/resources/common-english-words.txt, along with the addition of some common contractions.

followed by a comma, colon, or a hyphen, was b) used sentence finally and preceded by one of these, or c) began with an @ symbol (in which case the relevant string was everything after the @). "Sentence initial" was crudely operationalized as the first word in a post or a word following a period. Intuitively, words with a relatively high value for this should mostly be names for participants in the thread, sentence-modifying adverbials, and interjections. A major difference between these is that names, but not adverbials, will often be capitalized when not sentence initial, (interjections are expected to be intermediate because of often being put into all capital letters). Intuitively, words where a relatively high percentage of non-sentence-initial uses are capitalized should mostly be proper nouns, and probably only a minority of these will refer to users.

For this algorithm, I only included words appearing at least 10 times non-sentence-initially. For each word such word, the percentage of non-sentence-initial capitalized uses was calculated: The metric for name identification was (% non-sentence-initial-capitalized) * (%sentence-initial-vocative + %sentence-final-vocative + %beginning-with-at-sign), where each percent ranges from 0 to 100. A threshold of 1000 was used (below this point is when interjections began to appear), and all words at or above this threshold were treated as referring to users.

A.3 Running the classifier with and without removing names for users

In addition to these two algorithms, some additional names for users were identified by hand, as were some AVEN-specific terms (see Section 5.2.3 for details). To investigate the extent to which retaining these vs. removing them affected the accuracy of the classifier, I ran 160 trials of the classifier with these removed (discussed in Chapter 5.2.1), and also 160 trials without removing these.

Removing the names for users and AVEN-specific terms decreased the true positive rate, but only slightly: From 90.67% to 89.09%. Likewise, the true negative rate decreased, but

only slightly: From 83.96% to 82.16%. To determine whether to use unpaired t-tests or a Shapiro-Wilk tests, I tested for normality in all four lists of means {true positive rate, true negative rate} × {names retained, names removed} using the Shapiro-Wilk test for normality. For the Shapiro-Wilk test, the null hypothesis is that the data do come from a normal distribution. Thus, if the p-value is less than whatever alpha level is used, there is sufficient evidence that the data come from a non-normal distribution. I have decided to assume data are normally distributed if $p > 0.1$. (i.e. if the evidence against the null hypothesis is not even marginally significant.) The p-values ranged from 0.286 to 0.853. I thus assumed normal distribution for all cases, and ran an unpaired t-test for the asexual results in the two conditions ($p < 10^{-8}$), and likewise for the xkcd data ($p < 10^{-8}$).

To further understand the effects of removing names for users and AVEN-specific terms, Table A.1 compares the accuracy for both conditions for each forum in the AVEN-core, and Table A.2 shows results for each forum in the AVEN-other corpus (except the Site Info Center because of its size) and for each part of the asexual-other corpus. If removing (words identified as) names for users and AVEN-specific terms actually does what it is intended to do, we would generally expect a decrease in accuracy for the AVEN-core and AVEN-other forums, but little-to-no decrease in accuracy for the asexual-other data. Further, we would expect particularly large decreases in accuracy for forums that often talk about the site itself and/or specific members (e.g. Admods Archive, Site Comments).

For each forum, Pair-wise tests were run (classifier accuracy when removing vs. retaining names for users) to establish whether there was a statistically significant difference in accuracy. In order to determine whether parametric tests or non-parametric statistical tests should be used, Shaprio-Wilk tests were run on each forum for each condition. If p>0.1 for both, an unpaired T-test was performed. If p<0.1 or at least one of the conditions, a Mann-Whitney U test was performed.

The difference between the conditions with and without removing names for users is noticeably higher for the peripheral AVEN forums than for the non-AVEN asexual communities. This is consistent with the assumption that these terms would be more predictive

144

Forum	w/ names	w/o names	Difference	(test) p-value
Announcements	84.4015 (0.598)	76.0819 (0.6445)	8.3196	T: $p < 10^{-8}$
Welcome Lounge	97.9564 (0.0952)	97.4737 (0.0961)	0.4827	T: $p < 10^{-8}$
Asexual Q&A	95.3085 (0.1152)	94.9657 (0.1393)	0.3428	T: $p < 10^{-8}$
Asexual Relationships	96.4454 (0.1319)	96.0717 (0.1237)	0.3736	MW: $p < 10^{-8}$
Asexual Musings and Rantings	89.6148 (0.1425)	88.8713 (0.1378)	0.7435	T: $p < 10^{-8}$
Older Asexuals	79.0013 (0.4215)	75.6505 (0.4561)	3.3508	T: $p < 10^{-8}$
For Sexual Partners Friends and Allies	97.2476 (0.1952)	97.0543 (0.1913)	0.1933	T: $p < 10^{-8}$
Visibility and Education Projects	85.0912 (0.6228)	78.9703 (0.7343)	6.1209	T: $p < 10^{-8}$
World Watch	85.1845 (0.5923)	82.8578 (0.6559)	2.3267	T: $p < 10^{-8}$
Census Forum	83.3872 (0.4383)	82.6415 (0.4576)	0.7457	MW: $p < 10^{-8}$
The Gray Area	97.9469 (0.2586)	97.8934 (0.2747)	0.0535	T: p=0.075
Romantic and Aromantic Orientations	97.9892 (0.2563)	97.9638 (0.2478)	0.0254	T: p=0.37
Gender Discussion	82.7092 (0.3544)	81.4072 (0.3519)	1.302	MW: $p < 10^{-8}$
Celebration Time	78.2678 (0.5839)	75.2899 (0.5837)	2.978	T: $p < 10^{-8}$
Site Comments	78.6656 (0.5145)	66.1638 (0.6177)	12.5018	T: $p < 10^{-8}$
Admod Archives	83.4741 (0.5852)	74.1688 (0.6899)	9.3053	T: $p < 10^{-8}$

Table A.1: Mean (SD) accuracy for 160 trials of the NB classifier with and without names for users, by forum. "T indicates a t-test for independent samples. "WM indicates a Mann-Whitney U test.

Forum	w/ names	w/o names	Difference	(test) p-value
Just For Fun	52.54 (0.11)	45.28 (0.11)	7.27	MW: $p < 10^{-8}$
Off-A	55.59 (0.08)	53.24 (0.08)	2.35	T: $p < 10^{-8}$
Open Mic	65.2 (0.11)	62.0 (0.11)	3.2	T: $p < 10^{-8}$
Philosophy Politics and Science	29.5 (0.04)	28.17 (0.04)	1.33	MW: $p < 10^{-8}$
Worldpride 2012	67.56 (0.32)	62.48 (0.3)	5.07	MW: $p < 10^{-8}$
Apositive	87.51 (0.09)	85.29 (0.1)	2.21	MW: $p < 10^{-8}$
Blogs	82.67 (0.08)	82.07 (0.07)	0.6	T: $p < 10^{-8}$
HHA	86.27 (0.15)	85.33 (0.19)	0.95	MW: $p < 10^{-8}$
LJAsexuality	93.04 (0.03)	92.95 (0.03)	0.09	MW: $p < 10^{-8}$
LJAsexuals	91.54 (0.15)	91.9 (0.18)	-0.36	MW: $p < 10^{-8}$

Table A.2: Mean (STD) accuracy for 160 trials of the classifier applied to the AVEN-other forums and to other asexual communities, with and without removing names for users and AVEN-specific terms. "T indicates a t-test for independant samples. "WM indicates a Mann-Whitney U test.

of AVEN discourse than asexual discourse, and bias the classifier somewhat towards AVEN discourse. Removing them does have a statistically significant effect in the classifier's accuracy for all cases (although for one non-AVEN community, this was actually improved accuracy), but these seem quite small compared to the differences for the peripheral AVEN forums, although they are comparable to the differences for some of the core AVEN forums.

Appendix B

Dispersion metrics

Most keyness metrics—including the one that I use—are based on the assumption that the probability of any word appearing in a corpus is independent of any other word appearing (i.e. a bag of words model). It is well-known that there are a multitude of ways that this assumption is wrong. Indeed, a number of metrics used for studying language are based on the extent to which this assumption is violated (e.g. mutual information). From the perspective of making a list of keywords, the most important violation is that words tend to be "bursty". As Church and Gale (1995, p. 1, 8) explain, it "is well known that word rates depend on many factors: genre, author, topic, etc." and later, "Under standard independence assumptions, it is extremely unlikely that lightning would strike twice (or half a dozen times) in the same document. But text is more like a contagious disease than lightning. If we see one instance of a contagious disease such as tuberculosis in a city, then we would not be surprised to find quite of few more".

When examining the distribution of words among numerous documents, most words will probably be more bursty than predicted by a bag of words model, but some are more bursty than others. The relationship between burstyness and keyness is somewhat complex. Words that are more bursty tend to be more ones with more semantic content, such as proper nouns and technical terms, whereas the words that tend to be closer to what would be predicted under the bag of words model are the sort that are commonly used in a host of topics (Church and Gale, 1995). In information retrieval, this idea is used for a commonly used metric for identifying topics: TF.IDF (term frequency-inverse document frequency) is based on the idea that the important words for identifying the topic of a document tend to

be ones that a) have relatively high frequency in that document, and b) occur in a relatively low number of documents.

By contrast, when trying to identify important differences between relatively large collections of text consisting of many documents (e.g. comparing two corpora), words strongly concentrated in one particular part of the corpus are an easy source of false-positive results, especially when metrics are used that assume a bag of words model. For example, in the AVEN-core corpus *nad* (usually a misspelling of *and*) occurs 3855 times, but only and 17 times in the control. Also, *hic* (sound for a hiccup) occurs in the AVEN-core 905 times and in the control 21 times. Using plus-ten smoothing and removing names for users, these are ranked as the 22nd and 100th highest keywords. Closer analysis reveals that 95.7% of the uses of *hic* come from a single post containing 866 repetitions of "HIC!" and about $\frac{2}{3}$ of the remaining uses are from the same user. Although the uses of *nad* are not as concentrated in any one post (max = 5) or thread (max = 11), 97.12% come from a single user with a propensity for certain variant spellings—and who has the largest number of posts in my AVEN data.

In the context of linguistics, measures of dispersion (or of burstiness) generally take one of two general approaches. The first treats all of the text as a linear sequence of words (i.e. the first word is followed by the second word, which is followed by the third word, etc.) and measures the distances between these (or shows a picture). The second approach involves using some means of partitioning the data, and doing calculations based on frequency counts.

The first kind works well within a single document, and visual displays of this are a basic tool of many sorts of corpus analysis. For example, in the book for the Python library NLTK(Natural Language Toolkit Bird et al., 2009), the very first language-derived plot is a dispersion plot. However, such approaches are more problematic when there are numerous divisions in the data without any obvious means of ordering them (e.g. the corpus uses language from numerous documents), and they require a much higher word frequency.[1] For

---

[1] For example, Altmann et al. (2009) and Altmann et al. (2011) both engage in dispersion-based analyses of the same set of data (about 200,000 million words from a Usenet group). The latter uses a parts-based metric, and limits analysis to words occurring at least 5 times, while the former uses a distance-based metric and limits analysis to words occurring at least 10,000 times.

the second kind, there are a great many ways that posts can be partitioned into groups, but the most important for present purposes are by user and by thread.

Gries (2008) reviews several metrics of dispersion that have been used in corpus linguistics, arguing that information about dispersion should be reported more frequently. He highlights various problems with existing metrics, arguing for the importance of developing better metrics that overcome these problems, and proposing one possible metric. One major problem with many measures is that they require the corpus parts to be equally sized (and then they rely on word frequencies). One way around this is to normalize the data (i.e. use relative frequencies rather than absolute frequencies). However, this works rather poorly for two kinds of cases, given that the size of the corpus parts (when partitioned by users) varies widely. First, if a word is used by someone with a very low word count, this can greatly exaggerate how clumped it is. Second, if a word is primarily used by a single user with a very large word count, this can greatly mask the extent to which it is clumped. As shown in Figure B.1, the number of words per user (and per thread) varies widely—each spanning five to six orders of magnitude. The distribution for words per user can be explained as being a consequence of two facts: 1) The number of posts per user has a roughly zipfian distribution (including a long tail of users with a single post), and 2) The number of words per user is moderately correlated with the number of posts per user.

Altmann et al. (2011) provides a dispersion metric that works much better for data where there is a large amount of variation in the parts of the corpus, although it has not previously been used in keyword research to exclude extremely clumped words. As in the present study, Altmann et al.'s data come from an online message board (Usenet). The basic idea for their metric (for users) is to compare the observed number of users who used a word with the expected number of users under a bag of words model (but holding constant the number words used by each user and the number of times each word appears). The comparison they use is $\frac{Expected}{Observed}$. The challenge is calculating the expected value. Let N_A be the total number of words (tokens) in the corpus. For a given word w, N_w is its frequency in the corpus, and for a given user i, m_i is the number of times that the user i used the word w. Let $\tilde{U}_i$ be the
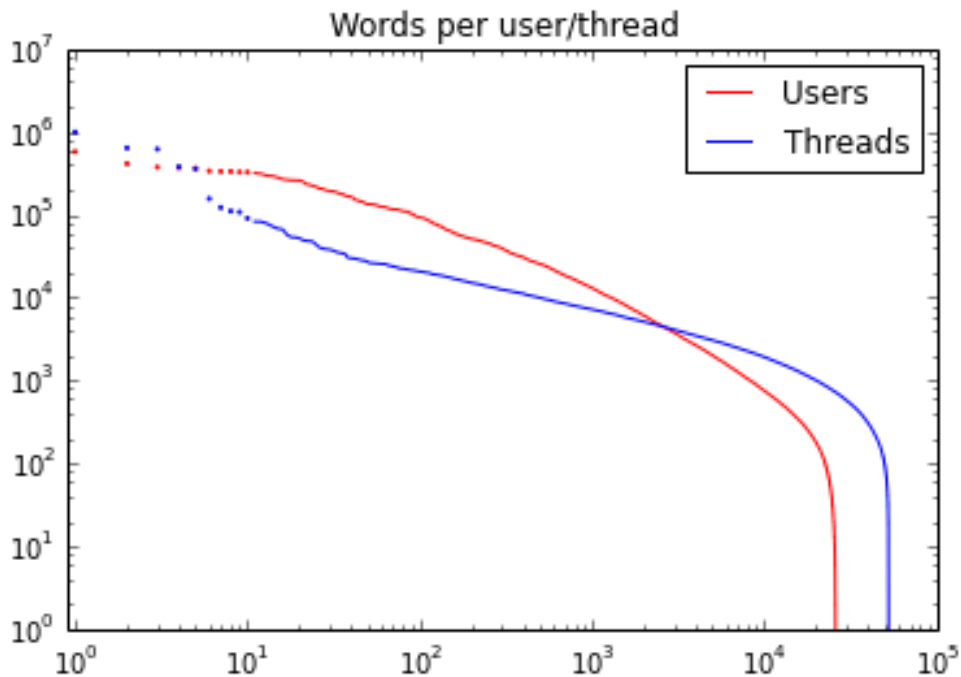
Figure B.1: Distribution of words per user and per thread in the AVEN-core corpus in rank-order.

probability that user i uses word w.

$$\tilde{U}_i = 1 - \prod_{j=0}^{N_w - 1} \left(1 - \frac{m_i}{N_A - j}\right) \approx 1 - e^{-f_w m_i}$$

where $f_w = \frac{N_w}{N_A}$, and the approximation is valid for f_w is much less than 1.[2] The expected value for a word, when partitioning the data by users, is the sum of the expected values for each user. The calculations for dispersion over threads is similar.

An additional benefit of this metric of dispersion is that, by making the value 1 to be chance, $\tilde{U} < 1$ means that the word is more clumped (according to users) than expected by chance, $\tilde{U} \approx 1$ means that a word is about as expected by chance, and $\tilde{U} > 1$ means that a word is more spread out than expected by chance. None of the metrics in discussed in Gries (2008) have this property.

---

[2]If this metric for dispersion were more widely used in corpus linguistics, it would need to be implemented in a corpus-linguistic software package, it would need to run reasonably fast, and this approximation dramatically reduces run-time. Another helpful optimization is based on the fact that it need not be calculated for each user and each word (above a certain frequency)–it only needs to be calculated for each token-frequency and each user-word-count.

Appendix C

Keywords

Table C.1: Top 200 keywords for asexual discourse, along with the uses per million words in the study corpus and in the control corpus, and the keyness score. Keyness is shown as the log_{10} of the ratio of the relative frequencies, after plus-100 smoothing.

	keyword	asex rfreq	ref rfreq	keyness value
31	aromantics	29.49	0.03	1.33
32	celibacy	52.35	1.12	1.33
33	homoromantic	29.1	0.04	1.32
34	orientations	57.62	1.51	1.31
35	cuddle	100.79	3.69	1.31
36	sensual	40.11	0.65	1.31
37	biromantic	27.17	0.01	1.3
38	aven's	27.01	0.01	1.3
39	bloomer	29.95	0.17	1.3
40	panromantic	27.66	0.07	1.29
41	sexuality	543.88	27.76	1.28
42	demisexuality	25.43	0.04	1.27
43	non-sexual	58.08	1.87	1.27
44	masturbate	104.67	4.5	1.26
45	attracted	536.2	28.79	1.26
46	friendship	219.17	11.51	1.24
47	closeness	39.89	0.99	1.24

Table C.1 (continued)

	keyword	asex rfreq	ref rfreq	keyness value
48	visibility	75.97	3.21	1.23
49	romance	200.8	10.85	1.23
50	sexually	441.65	25.63	1.22
51	affectionate	46.01	1.65	1.2
52	orientation	370.65	22.93	1.19
53	nonsexual	27.02	0.46	1.19
54	affection	89.84	4.73	1.18
55	platonic	87.15	4.6	1.18
56	neutrois	23.29	0.24	1.18
57	sexless	24.51	0.33	1.18
58	friendships	80.66	4.44	1.16
59	intimate	108.54	6.44	1.16
60	attractions	36.81	1.34	1.15
61	relationship	1629.56	117.69	1.15
62	androgynous	67.22	3.61	1.14
63	hetero-romantic	18.19	0.03	1.14
64	antisexual	17.98	0.01	1.14
65	squish	32.07	1.15	1.12
66	squishes	18.55	0.11	1.12
67	virgin	139.66	9.72	1.11
68	relationships	668.89	52.05	1.11
69	virginity	55.8	3.15	1.11
70	repulsion	24.15	0.62	1.1
71	fantasize	32.41	1.35	1.09
72	hetero	57.53	3.59	1.08

Table C.1 (continued)

	keyword	asex rfreq	ref rfreq	keyness value
73	flirting	69.8	4.64	1.08
74	partner	541.82	45.55	1.07
75	agender	22.8	0.66	1.07
76	aesthetically	44.66	2.62	1.07
77	anti-sexual	15.89	0.09	1.06
78	sex	4448.05	395.69	1.06
79	demiromantic	14.61	0.01	1.05
80	kissed	64.19	4.56	1.05
81	indifferent	60.49	4.25	1.05
82	partnered	19.66	0.52	1.04
83	cake	494.65	46.93	1.02
84	avenite	13.48	0.01	1.02
85	aromanticism	13.29	0.0	1.02
86	sexualities	23.77	1.05	1.02
87	off-a	13.44	0.03	1.02
88	demisexuals	13.73	0.06	1.01
89	companionship	31.45	1.85	1.01
90	genderless	25.65	1.34	1.0
91	kiss	161.42	15.31	1.0
92	flirt	33.94	2.21	1.0
93	cuddles	28.08	1.61	1.0
94	rantings	17.84	0.56	0.99
95	hugging	64.93	5.75	0.98
96	romantics	14.28	0.26	0.98
97	hi	421.05	44.57	0.97

Table C.1 (continued)

	keyword	asex rfreq	ref rfreq	keyness value
98	relate	186.05	19.07	0.97
99	sherlock	26.75	1.68	0.97
100	tmi	34.13	2.53	0.96
101	boyfriends	39.5	3.13	0.96
102	fantasies	68.4	6.44	0.96
103	dating	292.62	31.88	0.96
104	disinterest	15.97	0.53	0.96
105	crush	165.17	17.64	0.95
106	intercourse	74.47	7.3	0.95
107	aesthetic	67.47	6.5	0.95
108	feelings	513.23	57.95	0.95
109	bi-romantic	11.27	0.03	0.95
110	fetishes	29.1	2.18	0.94
111	masturbation	88.88	9.31	0.93
112	introverted	30.62	2.41	0.93
113	hypersexual	11.46	0.11	0.93
114	kisses	43.03	3.94	0.93
115	welcomes	16.08	0.68	0.93
116	faq	29.38	2.33	0.92
117	desire	545.77	66.2	0.92
118	lesbian	165.29	19.46	0.91
119	loner	17.47	0.96	0.91
120	hello	237.31	29.18	0.9
121	bi	92.16	10.58	0.9
122	sex-positive	15.4	0.72	0.9

Table C.1 (continued)

	keyword	asex rfreq	ref rfreq	keyness value
123	sex-repulsed	9.76	0.0	0.9
124	faqs	12.33	0.34	0.9
125	loneliness	23.03	1.74	0.9
126	orgasm	74.2	8.52	0.89
127	heteromantic	9.56	0.01	0.89
128	grossed	20.87	1.54	0.88
129	introvert	19.04	1.31	0.88
130	bisexual	86.91	10.55	0.88
131	crave	24.18	2.04	0.88
132	frigid	15.21	0.83	0.87
133	partners	150.71	19.52	0.87
134	personals	10.93	0.26	0.87
135	masturbating	35.18	3.61	0.87
136	repressed	34.23	3.58	0.86
137	congrats	66.98	8.28	0.86
138	foreplay	16.4	1.11	0.85
139	avenwiki	8.74	0.01	0.85
140	lonely	89.15	11.57	0.85
141	horny	39.32	4.46	0.85
142	husband	169.99	23.41	0.85
143	musings	18.82	1.49	0.85
144	tumblr	27.77	2.83	0.84
145	repulsive	29.56	3.1	0.84
146	boyfriend	291.51	41.7	0.84
147	therapist	60.87	7.79	0.84

Table C.1 (continued)

	keyword	asex rfreq	ref rfreq	keyness value
148	hyposexual	8.38	0.01	0.84
149	welcoming	24.11	2.4	0.83
150	fetish	53.21	6.88	0.83
151	acebook	8.21	0.03	0.83
152	homo-romantic	8.05	0.01	0.82
153	demis	8.12	0.03	0.82
154	genderqueer	26.11	2.85	0.82
155	pleasurable	22.97	2.37	0.82
156	androgyny	12.6	0.76	0.81
157	chat	127.64	18.96	0.81
158	virgins	24.09	2.62	0.81
159	nonlibidoist	7.66	0.0	0.81
160	arousing	15.34	1.25	0.8
161	disgusted	44.52	6.04	0.8
162	pansexual	19.8	2.01	0.8
163	heterosexual	125.09	19.2	0.8
164	schizoid	8.83	0.23	0.8
165	queer	85.37	12.78	0.8
166	desiring	15.61	1.37	0.79
167	tomboy	12.46	0.85	0.79
168	masturbated	11.66	0.72	0.79
169	ftm	15.13	1.29	0.79
170	dated	66.48	9.85	0.79
171	aspie	11.85	0.76	0.79
172	demi-sexual	7.18	0.0	0.78

Table C.1 (continued)

	keyword	asex rfreq	ref rfreq	keyness value
173	attractive	258.23	42.34	0.78
174	lgbtq	14.39	1.22	0.78
175	aro	7.16	0.0	0.78
176	grosses	9.49	0.42	0.78
177	partner's	20.5	2.31	0.78
178	greetings	19.06	2.07	0.78
179	relieved	29.87	3.94	0.78
180	autosexual	7.32	0.06	0.77
181	agendered	10.29	0.57	0.77
182	non-asexual	7.14	0.04	0.77
183	asexually	7.82	0.17	0.77
184	stimulation	28.51	3.81	0.77
185	montel	7.05	0.04	0.77
186	repulses	7.49	0.14	0.76
187	androgyne	8.61	0.34	0.76
188	felt	601.53	106.8	0.76
189	aspergers	16.96	1.87	0.75
190	urges	41.41	6.32	0.75
191	soulmate	10.03	0.63	0.75
192	non-romantic	7.39	0.16	0.75
193	amoeba	10.89	0.8	0.75
194	emotionally	115.27	19.95	0.75
195	pleasing	32.33	4.74	0.75
196	prude	12.14	1.05	0.75
197	labels	97.5	16.9	0.74

Table C.1 (continued)

	keyword	asex rfreq	ref rfreq	keyness value
198	label	197.47	35.65	0.74
199	fantasizing	11.04	0.88	0.74
200	erection	19.42	2.44	0.74

Appendix D

Items for Survey

1. Gender:
   Gender: Male
   Female
   Other (please specify)

2. Age
   (Given in 5 year intervals for 25-64, and 65+)

3. When did you first identity as asexual (using this specific term)? (If you do not remember the year, please guess as best you can)
   (dropdown menu for years)

4. Do you consider your asexuality to be a lack of sexuality?
   Yes
   No
   Other (please specify)

   5. In what country have you spent the most time in since age 12? (Choose only one)
   Australia
   Canada
   Ireland
   United Kingdom
   United States
   New Zealand
   Other (please specify)

6. What year did you first find an online asexual community?
   (dropdown menu for years)

7. What year did you first join an online asexual community?
   (dropdown menu for years)

8. How did you first find an online asexual community?

How did you first find an online asexual community? Entering "asexual" into a search engine

Entering a different search into a search engine

Media piece about asexuality (e.g. TV, news article, etc.)

Other (please specify)

9. Have you ever considered yourself "asexual"?

Yes

No

10. Did you ever use the term "asexual" to describe yourself before encountering online asexual communities?

Yes

No Other (please specify)

11. If you answered "yes" to the previous question, how/when did you use the term "asexual" prior to finding an online asexual community? (If you answered "no" to the previous question, skip this question.)

Only to myself

I only told one or a few people that I was asexual.

I told numerous people that I was asexual.

12. As a term applying to people, did you think of the term "asexual" on your own? (e.g. If you first heard the term as a means of reproduction, and then created, on your own, a new meaning of asexual that referred to people, please answer "yes.")

Yes

No

13. If "no," please indicate where you first heard "asexual" applied to people.

(write in)

14. In what sense(s) did you understand "asexual" when first considering yourself asexual? (Choose all that apply)

Not sexual.

Not sexually active.

On analogy with hetero/homo/bisexual

Other (please specify)

15. In what sense(s) do you presently understand "asexual"?
    Not sexual.
    Not sexually active.
    On analogy with hetero/homo/bisexual
    Other (please specify)

16. Have you ever considered yourself "nonsexual"?
    Yes
    No

17. Did you ever use the term "nonsexual" to describe yourself before encountering online asexual communities?
    Yes
    No
    Other (please specify)

18. If you answered "yes" to the previous question, how/when did you use the term "nonsexual" prior to finding an online asexual community? (If you answered "no" to the previous question, skip this question.)
    Only to myself
    I only told one or a few people that I was nonsexual.
    I told numerous people that I was nonsexual.

19. Do you presently use the term "nonsexual" to describe yourself?
    Yes
    No
    Other (please specify)

20. As a term applying to people, did you think of the term "nonsexual" on your own?
    Yes
    No

21. If "no," where did you first hear the term?

22. In what sense(s) did you understand "nonsexual" when first considering yourself nonsexual? (Choose all that apply)
    Not sexual.
    Not sexually active.
    On analogy with hetero/homo/bisexual
    Other (please specify)

23. In what sense(s) do you presently understand "nonsexual"?
    Not sexual.
    Not sexually active.
    On analogy with hetero/homo/bisexual
    Other (please specify)

24. Have you ever considered yourself "celibate"? (This is not asking about whether or not you are or were sexually active, but whether you have used the term "celibate" to describe yourself.)

    Yes
    No

25. Did you ever use the term "celibate" to describe yourself before encountering online asexual communities?

    Yes
    No
    Other (please specify)

26. If you answered "yes" to the previous question, how/when did you use the term "celibate" prior to finding an online asexual community? (If you answered "no" to the previous question, skip this question.)

    Only to myself
    I only told one or a few people that I was celibate.
    I told numerous people that I was celibate.

27. After finding an online asexual community, did you consider yourself celibate?

    Yes
    No
    Other (please specify)

28. Do you still consider yourself celibate?

    Yes
    No
    Other (please specify)

29. Have you ever considered yourself "abstinent"? (This is not asking about whether or not you are or were sexually active, but whether you have used the term "abstinent" to describe yourself.)

    Yes
    No

30. Did you ever use the term "abstinent" to describe yourself before encountering online asexual communities?

    Yes
    No
    Other (please specify)

31. If you answered "yes" to the previous question, how/when did you use the term "abstinent" prior to finding an online asexual community? (If you answered "no" to the previous question, skip this question.)

Only to myself

I only told one or a few people that I was abstinent.

I told numerous people that I was abstinent.

32. After finding an online asexual community, did you consider yourself "abstinent"?

Yes

No

Other (please specify)

33. Do you still consider yourself "abstinent"?

Yes

No

Other (please specify)

34. Prior to taking this survey, had you ever heard of Hypoactive Sexual Desire Disorder or Inhibited Sexual Desire? (If no, please go to the next page.)

Yes

No

35. Prior to finding an online asexual community, had you ever heard of Hypoactive Sexual Desire Disorder or Inhibited Sexual Desire? (If no, please go to the next page.)

Yes

No

36. Where did you hear about Hypoactive Sexual Desire Disorder or Inhibited Sexual Desire?

(write in)

37. Were you ever diagnosed with Hypoactive Sexual Desire Disorder or Inhibited Sexual Desire by a health/mental health professional?

Yes

No

Other (please specify)

38. Did you consider self-diagnosing? (That is, did you read about to try to decide if you have Hypoactive Sexual Desire Disorder or Inhibited Sexual Desire?)

Yes

No

Other (please specify)

39. Did you self-diagnose yourself with it? (That is, did you decide that you did had Hypoactive Sexual Desire Disorder or Inhibited Sexual Desire?)
    Yes
    No
    Other (please specify)

40. Prior to taking this survey, had you ever heard of Sexual Aversion Disorder? (If no, please go to the next page.)
    Yes
    No

41. Prior to finding an online asexual community, had you ever heard of Sexual Aversion Disorder? (If no, please go to the next page.)
    Yes
    No

42. Where did you hear about Sexual Aversion Disorder?
    (write in)

43. Were you ever diagnosed with Sexual Aversion Disorder by a health/mental health professional?
    Yes
    No
    Other (please specify)

44. Did you consider self-diagnosing? (That is, did you read about Sexual Aversion Disorder to try to decide if you have it?)
    Yes
    No
    Other (please specify)

45. Did you self-diagnose yourself with it? (That is, did you decide that you did had Sexual Aversion Disorder?)
    Yes
    No
    Other (please specify)

46. Just like many communities with their own specific interests, asexuals have created several new terms. Prior to finding an online asexual community, were there any words that you created relevant to asexuality?
    Yes
    No

47. If you answered "yes" to the previous question, please list the words that you invented. If you answered "no" please proceed to the next page.
    (5 write in boxes)

48. For each of the words you listed (if any) please explain your intended meaning(s).
(write in box)

49. Prior to finding online asexual communities, did you participate in any offline asexual communities or organizations? (Ones that specifically used the term "asexual.")
Yes
No

50. If yes, what communities or organizations? (If you feel comfortable doing so, please indicate the name(s) of the communities or organizations and give a brief description. If you feel this may be personally identifiable information, then just give a brief description.)
(write in)

51. Prior to finding online asexual communities, did you participate on any communities or organizations for issues related to asexuality (e.g. anti-sexual, celibate, abstinent, etc.)? (Please exclude any communities you may have listed in the above question.)
Yes
No

52. If yes, what communities or organizations? (If you feel comfortable doing so, please indicate the name(s) of the communities or organizations and give a brief description. If you feel this may be personally identifiable information, then just give a brief description.)
(write in)

53. In addition to being asexual, are you gay, lesbian, bi, or transgender?
Yes
No
Other (please specify)

54. Prior to finding an online asexual community, did you consider yourself part of LGBT/queer?
Yes
No
Other (please specify)

55. Subsequent to finding an online asexual community, have you consider yourself part of LGBT/queer
Yes
No Other (please specify)

56. Prior to finding an online asexual community, have you participated in one or more offline LGBT/queer communities?
Yes
No
Other (please specify)

57. Subsequent to finding an online asexual community, have you participated in one or more offline LGBT/queer communities?

    Yes

    No

    Other (please specify)

58. Prior to finding an online asexual community, did you actively look for information on what you now call asexuality?

    Yes

    No

59. If "yes" please list up to 5 places where you looked for information and rate each from 1-5 on how helpful it was for understanding (what you now call) asexuality, with one being the least helpful and 5 being the most helpful. (Type the number on the same line, after entering where you looked for information.)

60. Before finishing this survey, is there anything you would like to say?

Appendix E

IRB approval letter for Survey

The survey reported in this dissertation received IRB approval.

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

School of Literatures, Cultures and Linguistics

Office of the Director
4072 Foreign Languages Building, MC-171
707 South Mathews Avenue
Urbana, IL 61801-3675
USA

12/14/2011

Marina Terkourafi
Linguistics
4080 FLB

Dear Marina:

I am pleased to inform you that your project, *"Asexual identity before and after the creation of online asexal communities,"* was approved on 12/14/2011, for use of human subjects. You may now begin the data collection process.

Please be aware that this protocol was approved as is, any amendments to the protocol must be submitted for review. You are responsible for knowledge contained in the UIUC handbook, "Handbook for Investigators: For the Protection of Human Subjects in Research." Any unusual events or problems with this protocol should be reported to the SLCL Human Subjects Review Committee.

This approval will expire on 12/14/2013.

Sincerely,

Melissa Bowles
Associate Professor, Spanish, Italian and Portuguese
Director, Second Language Acquisition and Teacher Education
Chair, SLCL Human Subjects Review Committee

xc: James Yoon
      A Hinderliter - ahinder2@