Harry Potter Fandom as an Online Speech Community
Therese Katharina Skatun
Bryn Mawr College

Abstract

This paper locates the analysis of ONLINE SPEECH COMMUNITIES in the wider context of speech communities and COMPUTER MEDIATED COMMUNICATION (CMC), supplementing the rather scarce analysis of online speech communities. It uses the experiences gained by Cassell and Tversky (2005) and Wick (1997), and examines how their methods and findings are applicable to the larger and more open speech communities to be found online, specifically that of the Harry Potter fan community. Through the close examination of one Harry Potter fan forum, the Leaky Lounge, this study explores the ways in which members reinforce the community linguistically. Several markers were investigated, covering pronoun usage, different modes of member interaction, CMC proficiency, the use of community specific lexical markers, community norms, and the members' shared fantasy theme. The quantitative analysis established that most of the community reinforcing behavior increased as members gained in membership seniority, though some markers experienced a temporary decrease as novice members seek to prove themselves as *HP-sleuths*. The analysis also found that age and gender sometimes play a significant role in impacting online linguistic behavior. The results of the qualitative analysis likewise showed that members reinforce their community through conventional linguistic expressions such as greetings, but also through several methods particular to CMC, such as quotes, emoticons, and descriptions of their own actions. The latter were found to also play an

important role for the fantasy theme being acted out (virtually) by members in order to increase their range of expression and for humor.

1. Introduction and Background

Certain aspects of computer-mediated communication (CMC), especially the language of chat groups, have attracted attention of linguists (see e.g. Greenfield and Subrahmanyam 2003). This is probably due to its unique and highly marked features, dropping conventional rules of spelling, capitalization, and punctuation, and extensive use of abbreviations. The other aspects of computer mediated communication are also increasingly being discussed in the literature (see e.g. Crystal 2004).

Crystal's book *Language and the Internet* is the beginning of a more comprehensive look at the linguistic features of several online communication media, dividing it into five broad categories of Internet based communication that use language differently. These are electronic mail (e-mail), synchronous chat groups, asynchronous chat groups (news-groups and message boards), virtual worlds (MUDs), and world wide web. Crystal examines the linguistic conventions used in these separate media, and how they differ from not only real life speech and traditional forms of writing, but also how they differ from each other, recognizing that the language of chat groups is not the only 'genre' of the Internet. He also examines some of the principles underlying the various forms of CMC, which despite being different from speech and traditional forms of writing still have their own communicative coherence (Crystal 2004; 48).

However, not many studies have been done to look at online communication in the context of online speech communities. The notion of a speech community is a very

useful tool in sociolinguistic analysis, pioneered by such linguists as William Labov in the early 1960s. For example his 1961 Martha's Vineyard study of social attitudes underlying the use of island markers, moved beyond the study of isogloss boundaries to reveal how factors such as social attitudes and professional aspirations shaped language use (Milroy 1980;6-7). But while looking at factors such as class, age, sex, and profession, early sociolinguistic inquiry remained geographically bound in that they looked at linguistic differences within communities that shared both speech and geographic location. For example Leslie Milroy (1980) defines a community as "a local unit, characterized for its members by common locality and primary interaction" (Hymes in Milroy 1980;7).

However, this definition of community for linguistic purposes has since been changed to accommodate a wider conception of community, including ethnic and professional communities that exist in a wider area. Ronald Wardhaugh points out the inherent difficulty of defining a speech community, because it relies on how one defines both the group and a language variety (Wardhaugh 2002; Ch. 5). For a community to be a speech community, it must share norms of communication. With the concept of both community and speech variety being relative, it is therefore helpful to use a definition of speech community which allows for speech communities of differing scale: one may contrast Boston with New York, but one may also contrast the North East USA with the South, the Chicano speech communities with English and Spanish speech communities, and so forth (Wardhaugh 2002;121). In this regard it is also useful to recognize that these speech communities intersect; an individual will be part of more than one speech

community, defined by location (on several scales), profession, and so forth. As pointed out by Bolinger, (1975;333 in Wardhaugh 2002;124)

> "There is no limit to the ways in which human beings league themselves together for self-identification, security, gain, amusement, worship, or any of the other purposes that are held in common; consequently there is no limit to the number and variety of speech communities that are to be found in a society."

This move away from relying on location is both helpful and necessary in accounting for non-geographic speech communities. This understanding can then be applied to the study of online speech communities. Online communities are a new form of community that only surfaced in the 1990s, and they have been very little studied in terms of their behavior as speech communities. The notion of speech communities as intersecting is crucial to the study of online speech communities, because it is impossible to conceive of an individual only belonging to an online speech community. An individual who participates in an online community is at the same time a member of several real life (as opposed to virtual) communities, which generally dominate their participation in online speech communities in terms of both time and linguistic interaction.

The fact of online communities is becoming increasingly real. It is estimated that more people use the internet to participate in virtual communities than to make purchases (Horrigan 2001 in Porter 2004) and that eighty-four percent of Internet users have been involved with an online community (Bressler and Grantham 2000 in Porter 2004;2). Though the growth of virtual communities along with the growth of the Internet is clear, there is little consensus on how to define *virtual community* (Komito in Porter 2004;2).

Porter suggests a broad definition of virtual community that is useful for the purpose of studying online speech communities:

> "A virtual community is defined herein as an aggregation of individuals or business partners who interact around a shared interest, where the interaction is at least partially supported and/or mediated by technology and guided by some protocols or norms." (Porter 2004;3)

A narrower definition of online communities is proposed by Jones, who argues that online communities must meet the requirements of a true community. The four conditions for this are: "(1) a virtual common-public-space where a significant portion of interactive group-CMC occurs; (2) a variety of communicators; (3) a minimum level of sustained stable membership; and (4) a minimum level of interactivity" (Jones 1997). This definition is much narrower than that of Porter, eliminating business communities and groups without a specific 'location,' as not meeting his criteria.

The purpose here is not to go too deeply into the debate about what constitutes an online community and the related question of what constitutes a community. However, some points from this discussion are useful in trying to approximate what an online speech community is.

A working definition of online speech communities would thus need to satisfy the following five requirements:

(1) A network of more than one individuals (as opposed to organizations)

(2) Minimum level of sustained stable membership over time

(3) Interaction is linguistic in nature (as opposed to economic)

(4) Communication is interactive (as opposed to unidirectional)

(5) Members must primarily use CMC to communicate

(6) Share a set of norms for linguistic behavior

Everything except the emphasis on CMC could also be said to be true of a real life (RL) speech community. The medium of communication is the defining factor that makes online speech communities different from RL speech communities. But because of the nature of the Internet it is necessary to emphasize a few of the 'givens' of RL speech communities. The first of this is that there needs to be a minimum stable membership, which the rapidly changing nature of virtual groups does not necessarily guarantee. Also, it is important to emphasize that the communication taking place is not only linguistic in nature, but that it is interactive. This means that all the members of the community must be communicating with other members. Thus a typical online newspaper or a blog[1] would not be considered an online speech community, because the communication is only going one way.

Little work has been done in the effort to combine the analysis of online communication with that of online communities. Two relevant studies have been conducted. In 1997 Nancy Wick examined Pednet, as Usenet listserv for political walkers, in her dissertation for a degree in Speech Communication. Through the study of all the posts (925) on the listserv during the span of six months, Wick found that the listserv did constitute a speech community. In reaching this conclusion, Wick examined patterns in discourse, communicative norms, and culturally distinct forms of the community (Wick 1997:21-22). For patterns in discourse, Wick examined how the

---

[1] Also known as *web logs*, blogs are personal diaries or editorials through which a person can share his or her life and ideas with other people.

inclusive/exclusive use of *we* served to define the borders of the community (in this case to include bicyclists but not recreational walkers), and in seeking to establish communicative norms, she examined meta-communication (talk about talk) and rhetorical invocation of norms, such as 'be brief.' For culturally distinctive forms she examined how Pednet members created a 'pedestrian revenge fantasy' that became a lasting feature in the speech community (Wick 1997:37).

More recently, Cassell and Tversky have done a revealing study about the linguistic behavior of one online community. They studied the language of the Junior Summit '98, a project bringing 3062 children (ages 10-16) from 139 countries together through online communication to discuss and plan ways to make the world better through the use of technology (Cassell and Tversky 2005:1). Cassell and Tversky found that the changes in the children's linguistic behavior through the three months of the project moved toward a community through the use of linguistic markers. The study used an expanded version of Pennebaker's Linguistic Inquiry Word Count and a manual qualitative marking of the posts to track changes in the children's linguistic behavior (see e.g. Pennebaker and Stone 2003). Among the variables they checked were the relative frequency of *I*-statements, *we*-statements, apologies, answering each other's questions, and agreement. They found that *I*-statements decreased in favor of *we*-statements over time, and that apologies, answers, and other indicators of community interaction increased.

This thesis will examine how online Harry Potter fandom behaves as a speech community. It seeks to put the new notions of online community and online speech in a common spotlight to see what a sociolinguistic analysis of the web will reveal. The

working hypotheses of this project were largely modeled on those of Cassell and Tversky, namely that there are certain linguistic markers that indicate the 'salience' of the speech community, and that these markers will increase over time. It also sought to explore which other factors, in addition to the seniority of membership, could be seen as driving linguistic behavior online.

2. Overview of the Harry Potter Speech Community

2.1 Brief Background of Harry Potter Fandom

For those unfamiliar with the phenomenon, the Harry Potter (HP) fan community takes its inspiration and common point of interest from the *Harry Potter* series written by the British author J.K. Rowling. The first book, *Harry Potter and the Philosopher's Stone* (*Sorcerer's Stone* in the USA) was published in 1996. Since then, five more books out of the intended seven and four movie adaptations by Warner Brothers have been released, all of them instant bestsellers or blockbusters. The books have been translated into 63 languages and have sold over 300 million copies in over 200 countries (BBC News Oct 4 2005). Along with the immense popularity among children, but also adults, there has emerged a large and very active Harry Potter fan community, referred to as *HP fandom*. Inherent in the concept of fandom is that it is intended to be by fans, for fans. The primary form of interaction is through CMC, particularly the Internet, and both J.K. Rowling herself and Warner Brothers are coming to support and interact with HP fandom through the Internet as well.[2]

---

[2] J.K. Rowling says on her website, jkrowling.com, that she created it for this explicit purpose. Warner Brothers use harrypotter.com for promotion and information, and also work with fan sites on questions of franchise (see e.g. <http://www.the-leaky-cauldron.org/MTarchives/week_2005_09_25.php#007626>).

2.2 An Overview of the mediums of communication used

What makes the HP speech community different from most previously studied online communities is that members do not rely on a single mode of communication. Most RL speech communities rely mostly on face to face spoken interaction, perhaps supplemented by phone conversations, written notes and letters, and some CMC such as e-mails. Likewise, the network of HP fans online is sustained through multiple modes of communication, however, most of them are written CMC.

2.2.1 Written CMC

Most of the mediums of communication used are the written electronic communication forms generally associated with online communication. The most visible of these are regular World Wide Web pages with information. They are always public, and appear at first sight to be unidirectional in that the staff post and then readers read. However, many sites now have a comments feature enabled, which links to another page where readers can give their reactions to staff members and to each other. Examples of such sites would be The Leaky Cauldron, Mugglenet, the Harry Potter Automatic News Aggregator, and the Harry Potter Lexicon. For links to these sites, please see Appendix 8.3.

Another, very different form of communication used frequently for private, interpersonal communication is electronic mail, in the form of e-mail or private messages within other kinds of software. There are also e-mail based groups that reach a larger

readership, but are still limited in their reach. E.g. Yahoo lists over 3000 of such Harry Potter groups, ranging in membership from 3000 to less then 5.[3]

In terms of synchronous chat, in which members are online at the same time, there are several different forms used for communication. There are traditional open chat groups about HP, even though they are not as central as other forms of communication.[4] Community members who meet through other kinds of communication will often continue their discussions and friendship through Instant Messenger programs. This communication is different from that of e.g. Internet Relay Chat (IRC) groups because it is intimate and less stressful. It is also different from e-mail and message boards because it happens in real time, or synchronously, so that an instant response can be expected. For example, all the about 40 staff members of The Leaky Cauldron (TLC) have Yahoo! Messenger and use it for both social and 'business' interaction.

Message boards (also known as discussion forums) are perhaps the place where one sees the online speech community most clearly. These forums are organized into thematic sections containing multiple (in some cases thousands) of discussions in the form of threads of posts organized chronologically. Like IRC they are on-going public discussions, but they differ in that they are asynchronous; the discussions do not go away, and one does not expect an immediate response. Also, posts tend to be considerably longer and more substantive than in synchronous chat, and adhere more to standard conventions of spelling and grammar, while still making use of chat elements such as emoticons (smileys). Examples of this include the Leaky Lounge Forum and the Chamber of Secrets Forums.

---

[3] http://groups.yahoo.com/search?query=Harry+Potter&submit=Search
[4] See e.g. Veritaserum chat room: http://www.veritaserum.com/chat/

Two related written forms of communication are fan works published online for reading by other community members. They fall into the categories of academic writing (at least comparatively academic in style) and fan fiction (creative writing). Most sites hosting essays/editorials and fan fiction encourage their discussion and provide the virtual means for readers to give their opinions on the pieces. Examples include Mugglenet Editorials, FictionAlley, SugarQuill.

The last form of written CMC surveyed here are virtual worlds (multi user dimensions - MUDs) and role playing games (RPGs), which are make believe worlds facilitated online. They are both entirely text based. MUDs rely on computer software and a series of written commands to go on adventures and interact with other users. Role playing games are the less high tech ancestor of MUDs, in which the game is controlled by a human.[5]

2.2.2 Spoken CMC

However, new technologies are making spoken CMC increasingly popular. While certain instant messengers have had the possibility of voice conversation built into their software, spoken conversation has become increasingly popular with Skype, a free software that allows for telephone conversations over the Internet. The use of this form of communication is still quite low compared to e-mail, but it is gaining momentum parallel with the release of PodCast technology by Apple. In August and September 2005, Mugglenet and The Leaky Cauldron each launched MuggleCast and PotterCast respectively, releasing about an hour of audio each week in the style of a radio show.

---

[5] For a more comprehensive discussion of MUDs and RPGs online, see Wikipedia: Harry Potter Fandom/RPGs.

Since the speech community has been (and still is) primarily sustained through written CMC, adjusting how to make use of this medium is still in progress. For example, in the first episodes of MuggleCast the speakers showed a lot of insecurity in how to pronounce certain words from the books, e.g. *Accio, Karkaroff.* Though certain lexical markers (e.g. *ship*, discussed later) do feature in the conversation, many of the other makers of the online speech community have taken time to become integrated, such as the use of shorthand abbreviations (e.g. *HP* for *Harry Potter*).

While Crystal focuses to a large degree on the limits of written CMC compared to spoken CMC (2004:42-43), some of the problems and linguistic debates encountered in the PodCasts indicate that written CMC enables linguistic expressions not available in spoken language. The most obvious example of this is punctuation and the non-letter characters of a standard keyboard. Two examples of such usage can be seen in section 4.6.3 on constructions, below. For example, *H/H* (indicating a romantic pairing between two characters whose names begin with H) is not distinguished in speaking from the abbreviation *HH* (usually indicative of Helga Hufflepuff). Yet instead of saying *Harry in a romantic relationship with Hermione*, or *Harry Potter*, there has been a gradual adoption of fandom short-hand in the spoken language of the PodCasts, so that people say the conventional written abbreviations *HH* (despite the ambiguity) and *HP*. Anecdotally, the dependence on written CMC has led HP fandom members meeting for the first time to continue to communicate via instant messenger while sitting in the same room.

2.2.3 Traditional forms of writing

The impact of traditional forms of writing in the HP online speech community have been minimal. With only a few exceptions the published literature on Harry Potter has a surprisingly low contact with the fan base. The literature tends to be either redundant for well-versed fans (e.g. *Kronzek and Kronzek, Colbert*) because it seeks to explain the obvious, or too dense or academic to interest the younger majority (e.g. *Granger*). One exception to this would be *The Plot Thickens*, a compilation of essays written by fans and edited by Galadriel Waters, who has published her own books on *Harry Potter*, and also uses the web to stay in contact with her readership (Waters 2004).

2.2.4 Real life spoken communication

Face to face encounters are not common, but they do happen at various conferences and events dedicated to the books, such as *Witching Hour*, a conference in Salem, Massachusetts, in October 2005. During such events a few hundred fans will spend some very intense days together, and then part. However, both the number and frequency of such events are very small compared to the amount of CMC that takes place.

All these forms of communication require skill to master, and particularly those of written CMC. Here they have been briefly outlined, and it is important to keep in mind that these CMC technologies vary in several respects, and do not always follow the same principles or allow for similar expressions. They vary in how public or private they are, on what time frame they operate, how interactive they allow 'speakers' to be, the range of

graphical flexibility (prosody) they allow, revisability, and expectations as to appropriate style and content.

2.3 Fandom Demographics

Expressing and tracking identity online

The collection of data about online demographics is difficult because the Internet is a place where many prefer to be anonymous (Thurlow 2004;99), and reliable data is still more difficult to procure. However, users of the internet do have several ways of expressing who they are that differ from RL obligatory markers such as race, style of dressing, location, age, and gender. One of the major features is username, the name under which a person communicates. Other features that depend on the CMC medium used are signature lines (a short piece of text or a graphic), avatars (image that appears next to your name), and profile pages, where one can give considerable amounts of information about oneself if one chooses. For an example of what an online 'face' looks like, see Appendix 8.1.

2.3.1 Traditional demographics

Sociolinguistic analysis has tended to look at how factors such as socio-economic class, locality, age, gender, ethnicity, and social aspirations influence our use of language. However, many of these identity factors do not have the same degree of importance in online speech communities. For some of these demographics it is not possible to find information. Below is a brief discussion of some of the demographic features of the HP community:

Though there is no data available, it is probable that the whole socio-economic spectrum is represented in the HP online speech community. There might be a predominance of people living in households wealthy enough to own a computer, but the existence of computers in schools and public libraries means that even poor dedicated HP fans in many countries would have access. In the community itself, this feature is entirely unmarked, and discussions of personal wealth (or lack thereof) is rare.

As with socio-economic class, ethnicity is an invisible feature online, and aside from cultural sharing, it is rarely mentioned. The traditional meaning of *race* (e.g. Caucasian, East Asian, Black) is a largely irrelevant factor for the HP online speech community not only because it is invisible, but also because of the thematic attitude of this community; the question of human race takes second place to the question of magical blood status and non-human races. The extent of this can be seen in the rules of the TLC forum, which explicitly bans "insults or harassment of others based on actual or perceived gender, sexual orientation, age, race, ethnicity, religion, magical status or lack thereof, Hogwarts house, national origin, political affiliation, character preference, or "ship" preference" (TLC: Rules). This fantasy theme will be discussed further in section 4.8.

The membership of the HP speech community is as international as the HP phenomenon itself. This information is technically recordable through the tracking of members' Internet Protocol addresses, but is not available. For example the Leaky Lounge forum has members from at least 35 different countries.[6] Following this wide

---

[6] This estimate was concluded by reading through the different topics of the forum's sub-forum for cultural discussions and counting the number of country threads.
<http://www.leakylounge.com/cgi-bin/ikonboard.cgi?s=7ca6305054e7aca2bf10d9c406797bc7;act=SF;f=53;st=0>

dispersion of members, it is clear that a large portion of the members do not have English as a native language.

The two demographics more easily available are age and gender, though the data is scattered and always self-reported. In a recent poll at TLC, a total of 50 000 polled members showed that the readership on the main page ranged from under ten to over sixty, but with a majority between the ages of fifteen and thirty. Age is more transparent than other demographics, because it is considered less personal, and because attitudes and even minimal personal information are very indicative (e.g. someone mentioning homework is probably younger than someone referring to their son's homework).

Like age, gender is more visible, and they are both commonly requested when joining a community. They can be discovered through a little effort by a fellow member (e.g. on a member profile kept by the software system or by asking). Questions about gender are not considered rude. When given, a member's real name is generally indicative of his or her sex. The more commonly used screen names are sometimes indicative, but not always.[7] Names can also be misleading, since people will assume the names of characters they like regardless of gender, or may wish to pass for a member of the opposite sex.

Although less so than in real life, the demographics that are most visible and accessible online are thus gender and age. The results found from looking at the linguistic behavior of members at the Leaky Lounge forum indicate that these two categories do matter. While some parts of language use were not gender-related, particularly the interactive aspect of communication showed differences in behavior

---

[7] Some names contain elements with clear gender indications, such as *girl, lady, Mrs., Miss* and *Mister, Mr., boy, man*. For example *paperflowergirl* (f), *Mr. Quint* (m). However, many names are not clear, e.g. *Puffin* (f).

between males and females similar to those found in the real world. Age also impacted linguistic behavior.

2.3.2 Online Demographics

When joining an online community (usually by registering a username and password and then joining the conversation), one is not automatically a 'full' member. Online communities have names for different seniorities of membership, progressing from *newbie*, to *novice*, and to *wizard*. The term *newbie* is very universal, and denotes newcomers in an online community, who may not be expected to fully grasp the technology, make blunders with the rules, and do generally 'silly' things in the eyes of others. Despite its applicability, *wizard* is not much used in HP fandom, perhaps because of its potential for confusion, and *experienced member* or *old-timer* is used instead. These are the members who have mastered the technical aspects of communication, know the rules, and may be somehow involved in the administrative aspects of the community, for example as moderators and technical staff. Novices are those in-between newbie and wizard. The importance of this distinction will be seen in the results, as members show changing patterns of communication as they progress from being newbies, to novices, and wizards.

2.3.3 Harry Potter Demographics

While many of the traditional identity markers are not very visible in the HP online speech community, there are new ones that are highly advertised. These are 'chosen demographics' in that they are directly related to the books and the individual

member's attitude to the books. The most common of these are favorite character, what romantic relationship a person supports (in fandom called *ship preference*), Hogwarts house, personal position on Snape, endorsement of speculative theories, and fascination with the Dark Arts. These are commonly expressed in user names (for example *Werewolf's Lover*), signatures ("Remember Cedric Diggory"), avatars (a picture of two characters hugging), and color use (using green text to denote Slytherin). The extent to which these opinions, sometimes even adopted identities, impact linguistic behavior is difficult to ascertain, and beyond the fact that those interested in a certain aspect of the books used more vocabulary related to this topic, no relationships were shown.

2.4 The Scope and Structure of Harry Potter Fandom

In looking at how 'salient' HP fandom is as a linguistic community, there are a few factors that should be considered: size, the density and nature of the network, who the members are, and the communicative activity of members.

2.4.1 Size

Defining the size of the community is difficult, because reliable data for the entire community is difficult to procure, and also because it is difficult to clearly define the limits of the community. However, certain statistics should give at least an indication of the scope of this community. Below is an overview of the traffic experienced by the major HP websites.

Table 1: HP site traffic indicators[8]

Site	Purpose	Alexa Rank	Alexa Reach	Poll feedback
Mugglenet	news, multi	2,825	335	100 000
Veritaserum	news	10,226	124	40 000
TLC	News, multi	13,180	117	30 000
HPANA	News	19,683	83	17 000
HP Lexicon	Lexicon	32,451	37	-
HarryLatino	News, multi	73,468	22	20 000

This data is taken from Alexa, an online aggregate that measures different parameters of site traffic. All statistics are the daily average over three months. Alexa Rank indicates how much traffic a site gets compared to other sites. Mugglenet is thus the 2835th most popular site on the web. Reach is numbers of (distinct) users per million. Thus it does not give absolute numbers, but approximates the percentage of web-surfers who visit this particular site.

A less 'advanced' but more concrete indicator of site visitors is to look at poll responses. Consider a poll on e.g. one's opinion on the sound track for the Goblet of Fire film: the total vote of 11000 indicates that at least this many users visited the site in the one week that particular poll was up (www.the-leaky-cauldron.org, October 27 2005). Polls are indicative, but also problematic, because visitors may choose not to vote at all, or may vote more than once. The polls may also be left up for varying lengths of time. Unfortunately, most sites do not archive voting turnout and time frame, so these numbers are very approximate.

This overview, however, is not entirely indicative of the size of the speech community (as opposed to fandom) because several of the people who visit only read and do not interact (they are lurkers). The exact numbers of lurkers is hard to determine. A greater approximation of actual online speech community membership can therefore be

---

[8] Alexa.com

derived from looking at forum membership statistics in Table 2 below. What these statistics do not show is that not everyone who registers ends up participating. For example, at the Leaky Lounge only about 4000 members have more than 10 posts, and it is probable that other forums, whose member lists are less accessible, have the same trend.[9] Even forums have their lurkers, and a random sampling of forum activity statistics over time show that about half the people visiting the Leaky Lounge at a given time are not registered members (who therefore do not post). For websites the percentage of lurkers is probably much higher.

Table 2:forums and membership

Site	Membership	Posts	Start date	Language
CoSForums	66 000	659 000	2002	English
HPANA forums	59 000	2 500 000	2003	English
Leaky Lounge	26 000	391 000	2005	English
Dark Mark	13 000	727 000	2002	English
HarryLatino	7 830	785 000	2003	Spanish

From the data in Table 1 and 2 it is clear that the membership in the HP community is quite large, and that the membership in the HP speech community, though smaller, is still several thousand people. This is clearly not the kind of intimate network structure (where everyone knows and communicates with everyone) some have deemed necessary for online communities (Jones 1997). However, size is not a limit for RL speech communities, and there is no reason to suppose that it should be so for online speech communities.

2.4.2 *Membership*

---

[9] See http://www.leakylounge.com/cgi-bin/ikonboard.cgi?act=Members

20

The question of membership is one that online communities themselves care about, and the following typology is one frequently used for this purpose (see Crystal 2004;52-53)

a. *Member*: individuals who actively participate in a community by both reading and responding (i.e. interacting) with other members of the community. For the purposes of a speech community, the members should further be defined as someone who participates and interacts on a fairly regular basis.

b. *Lurker*: individuals who read, even avidly, but who never make their presence known through responding or communicating with other members.

c. *Troll*: individuals who appear at first sight to be members of the community because they do make their presence known. However, these people participate in the community with the explicit purpose of subverting it through provocative and offensive behavior, and as such are not true members, and are eventually thrown out of the community if its technology permits it.

2.4.3 Network structure

In looking at speech communities, it is helpful to understand some of the network structure. The HP online speech community falls quite unequivocally into the category of a simplex community, as the members have only one point of interaction: the time they spend online discussing Harry Potter and other topics that emerge while they intend to talk about Harry Potter. It is a leisure activity, and very few people know and interact online with people they also meet in their real lives through work, kinship ties, geographic location, or social groups. The inherently anonymous nature of the Internet,

coupled perhaps with the fact that many are embarrassed to admit that they are "Potterheads" to RL colleagues, makes it impossible to show statistics on this. However, anecdotal evidence suggest that a lot of the people who come online to discuss Harry Potter do so because they do not have anyone in their real lives who shares their interest to the same extent.[10]

In terms of density, the HP online speech community is best described as a dense speech community, in that many of the community members are in contact with each other. Indeed, a lot of the communication is public, in the form of message boards and other forms of communication, and communication with only one or very few others is rare, almost impossible to achieve. However, there still are varying degrees of intensity in communication between the speakers, depending both on how much time the individual speaker is able to spend online, and on where in the network a person is located. Though HP online fandom can be described as grass roots, it still has an institutional structure that impacts how communication takes place.

The two ways of mapping the online speech community network are outlined below.

Figure 1: simplified network of speaker connectedness.

■ Staff and moderators

▨ Members

---
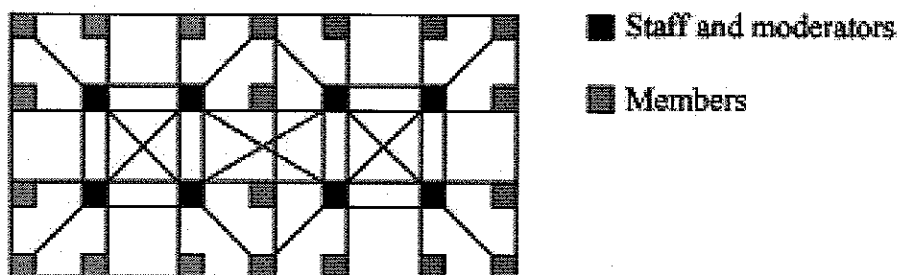
[10] For instance, the author's friends and family tend to stop listening after about five minutes of intense Harry Potter conversation, but at the Leaky Lounge, she may enjoy talking to people who are interested in listening and responding.

Figure 1 illustrates how the density of the speech community is centered around the various online sites dedicated to Harry Potter. Some members mainly participate on one site, while others participate in several places, and there is a substantial overlap in membership. This is shown by the fact that for all the larger Harry Potter sites listed by Alexa.com, the ten other most frequently visited sites by its users are also all Harry Potter sites.

However, the flow of communication within these sub-networks is not even among all the members, nor is it even in its direction. The staffs of these respective pages engage in a fair amount of mass, one-way communication, which reaches all the members, but not all the members' communication reaches the staff. An example of this would be the very popular news sites, dedicated to sharing updates of HP news. Nevertheless, the staff members are very involved in the communities, through answering e-mails, moderating forums, and approving fan fiction or essay submissions depending on their job. The communication between staff members is generally more frequent and denser than that between members. This is because the coordination of large sites and the decision making processes involved requires constant communication to make sure everything runs smoothly. The staffs of various sites also interact frequently on collaborative projects, such as joint PodCasts and various projects (e.g. PotterParties.com, Floo-Network.org). Being both potential collaborators and competitors, the staffs of various sites like to keep very up to date on what other sites are doing.

Figure 2: simplified network of HP websites
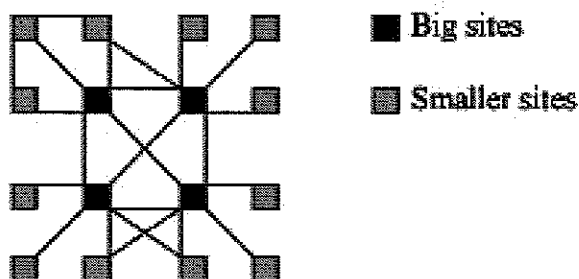
■ Big sites

▨ Smaller sites

Figure 2 outlines the network between various news sites. The traceable relationship between various sites is less dense than that between members. This is because there is a central group of very large, popular, and high quality English language sites on which the less popular sites rely for their content, which is mainly news. Thus, one would expect linguistic inventions to spread primarily through the staff of various HP sites, and particularly the staff of the large and popular sites to non-staff members and smaller sites.

2.4.4 The HP speech community in the context of other online speech communities

Since Harry Potter fandom exists primarily online, there is a considerable influence of the general linguistic trends of CMC also in the HP online speech community. It is probable that the members of the speech community will also use CMC technology to communicate with non-members, such as chatting with RL friends over instant messenger, private and professional e-mailing, e-commerce, online newspapers, and so forth. As described by Crystal (2004), all of these genres of CMC take distinct forms, and the general linguistic features of the internet, such as the use of numerous abbreviations, are also present in the HP speech community. This will be discussed in further detail in section 4.5.

There are also other online communities in which HP online speech community members are likely to participate. Among these are other fan communities, particularly those centered around fantasy and science fiction, such as *Lord of the Rings* fandom, and *Star Wars* fandom. The evidence for the popularity of other works of fantasy is mostly anecdotal, but is supported by the fact that a large part of the non-HP books, films, and TV programs discussed on HP forums fall into the science fiction and fantasy genre.

There is also an overlap in membership with the fan fiction community. For those unfamiliar with this phenomenon, fan fiction involves fans taking the characters, setting, or other part of an authors story and using them to create their own stories. There are sites dedicated to HP fan fiction only (such as FictionAlley), but a lot of HP fan fiction is also hosted on sites with a wide variety of fan fiction (such as Fanfiction.net).

Figure 3: Overlap between different online fandom membership
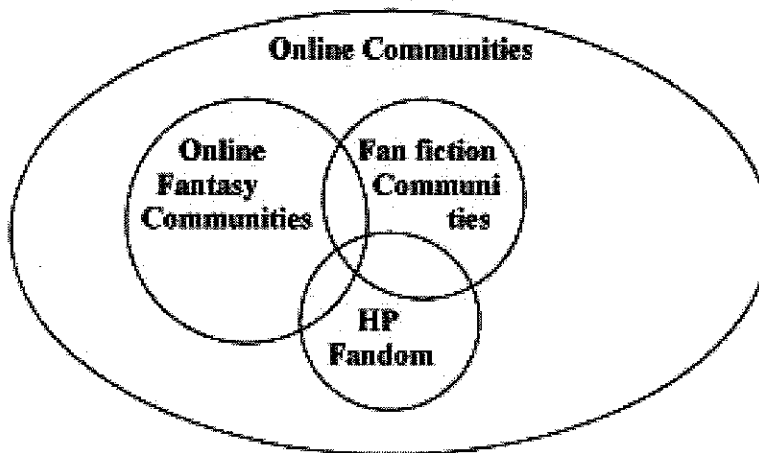
2.4.5 Sub-groups within HP fandom

Though various sub-groups do exist, it is hard to determine how solid they are. However, one can see groupings around particular sites, since it is likely that some members will chose one or two sites at which to participate more regularly than other

sites. This will particularly be the case in instances where language barriers constrain individuals to only participate in a certain language. Also those involved in writing, reading, and reviewing fan fiction belong to a more specialized community than broader community. It is possible that the various traditional or HP demographics discussed under those headings are strong enough to form cohesive sub-groups. However, no linguistic analysis has been done comparing language use within different sub-groups. The bounds of fandom linguistic markers (such as particular vocabulary and expressions) do extend beyond the English-speaking world. How and if they differ, aside from the obvious native language difference, would be an interesting line of inquiry, but it will not be pursued here.

3. Analyzing HP fandom as an online speech community

3.1 Rationale

Above I have discussed some of the network-structure and demographic features of HP fandom as an online speech community. The following section of this paper will look at the ways in which the members of the speech community linguistically reinforce their status as a community. As mentioned earlier, a very enlightening study to this effect was carried out by Cassell and Tversky (2005), and the basis for the following method has been adapted from their research on an online speech community that was limited in both time and size. The working hypothesis was that the results for the Leaky Lounge would be similar to those of the Junior Summit '98, namely that linguistic reinforcement of the community increases as members become more integrated.

Instead of a purely qualitative approach, such as the one taken by Wick (1997), the approach taken here has a strong quantitative aspect. The purpose behind the statistical analysis described in the following sections, was to better understand the formation of an online speech community through the members' linguistic behavior. The primary interest was thus in tracking changes in usage of verbal community re-enforcers over time. In an attempt to uncover what other variables, in addition to 'seniority' influence the linguistic behavior of the members of this online speech community, gender, age, and posting frequency were also examined.

The specific markers of the HP online speech community were also examined, and empirical analysis was conducted on the use of specialized vocabulary and abbreviations. In addition to this, other forms of markers, such as certain constructions, expressions, and socially conscious attitudes were examined qualitatively.

Keeping in mind the wide variety of CMC technologies available, the empirical analysis will take its base in the discussion board / message board technology, particularly that of the Leaky Lounge, the discussion forum of The Leaky Cauldron. This choice was made because as a staff member of TLC, I know the board well. The forum webmaster, Nick, was crucial for devising a way to accessing member posts in a usable format.

3.2 Methods

The data was mined from 15 members of the Leaky Lounge Forum. The members were selected through a largely random process from a limited pool of members. This pool was constrained by the following factors:

a. The member must have made over 500 posts. This was to get a sizeable sample of posts and to ensure that the Forum member was interacting enough with the Forum to be considered a part of the speech community.

b. The member must not have been a repeated rule-breaker. This is also to make sure that the Forum member is an earnest member of the community.

c. The member is not staff. A significant number of the high posting members at the Leaky Lounge end up becoming staff. Though they are the people who can be said to be most intensely involved in the community and the speech community, the tasks involved produce a large amount of certain kinds of communication, such as "I am closing this thread because a discussion on this topic already exists here. Have a magical day, [wizard emoticon]." One member in the sample is currently staff, but was not so during the first 1000 posts considered here.

d. Only members with both age and gender information in their profiles were included. Even if there is no guarantee that this information is correct, having this information is useful.

e. The last constraint placed on the selection of members is that it was divided into two pools: one female, one male. The majority of Leaky Lounge members are female, and a completely random sample would probably produce too few males to make gender comparisons possible. There are thus seven professed females, and six professed males in the sample.

Table 3: Overview of member sample

Member	Joined	Posts	Frequency (posts/day)	Age	Age group	Gender	Location
Member1	Jan	9980	35.4	20	Mid	f	UK
Member2	Feb	703	2.5	35	Old	f	Australia

Member3	Jan	1592	5.6	17	Young	f	Canada
Member4	June	982	7.3	28	Mid	f	US
Member5	June	537	3.8	35	Old	f	US
Member6	June	983	6.4	22	Mid	f	US
Member7	July	1380	12.2	18	Young	f	US
Member8	Jan	674	2.4	47	Old	f	US
Member9	Jan	1412	5	18	Young	m	US
Member10	Jan	627	2.2	25	Mid	m	Philippines
Member11	May	573	3.3	18	Young	m	Belgium
Member12	July	675	5.7	24	Mid	m	Mexico
Member13	June	857	5.6	21	Mid	m	Netherlands
Member14	July	809	7.5	29	Old	m	Canada
Member15	July	1012	9.8	24	Mid	m	US

All data in the table pertaining to age, gender, and location is self-reported, except for the fact that a few members' location was ascertained by finding the Internet Protocol address of their most recent post.[11] All members write fluent English with no noticeable mistakes, so it is difficult to ascertain who is a native English speaker, and who is not, though location is the closest proxy available. Direct inquiries of the members were not made because this would violate the rules of the Leaky Lounge forum (The Leaky Cauldron: Rules).

The data was organized by member, in strings of chronologically ordered posts, regardless of where (in which discussion) the post was made. The data covered by this size of sample amounted to approximately 9000 posts, since all but one of the sample speakers had more than 600 posts. Many had significantly more posts, but the extra data between 600 and 1000 was used sparingly. Beyond 1000 posts no data was collected. This data was then parsed by a set of programs that counted the occurrence of certain words, phrases, or html code, per post. The first program, Magic_grep, ensured that only the poster's own words were part of the linguistic sample, excluding the forum's html and

---

[11] Internet Protocol (IP) is the standard for identifying computers through a set of numbers, which also ascertain the geographic location of the server used by the member.

quoting of other members.[12] The frequency per post of certain variables, such as self-reference, apologies, and emoticons was found by writing out all the possible combinations of these, helped by some program features.

However, by seeking to match certain strings of characters to the occurrence of such in the data, there will obviously be many instances that could only be recorded through manual coding. This was part of the method employed by Cassell and Tversky (2004), and by Nancy Wick (1997). This was unfortunately not possible given time, resources, and the sheer size of the sample. Though frequent spelling variations and typos were taken into consideration, badly misspelled words were hard to predict and include in the search strings. Also, unusual expressions and implied messages were not accounted for. A fine balance was sought between searches that would capture all instances of a certain linguistic behavior, but would at the same time include only relevant linguistic behavior. For the full set of search strings used, please see Appendix 8.2.

Another program then aggregated the data into groups of twenty posts. The output per 20 posts was then compared to overall linguistic output (number of words) per twenty posts to produce a frequency score. This corrects for the fact that some members write considerably longer posts than others, unduly influencing the sample, and for the fact that the overall linguistic output did not stay constant over time. The frequency scores were then used to compute averages per twenty posts (1-20, 21-40, 41-60 and so forth) to track changes over time for each variable in a spreadsheet. Averages were also computed by gender, age, and posting frequency, in an attempt to uncover what other

---

[12] Credit for the brilliance of the programs I used belong to Sean Crist, who kindly wrote them for this project.

variables, in addition to 'seniority' influence the linguistic behavior of the members of this online speech community. Age computations were divided into three groups, as seen below. Some adjustments of the age-brackets were made to have as even a distribution of males and females as possible, though entirely even distribution was not possible due to the small number of the sample.

Table 4: Members by age group

Age group	Age range	Female	Male	Total
Young	13-19	2	2	4
Mid	20-28	3	4	7
Old	29-up	3	1	4

All calculations and graphs were made using Microsoft Excel. Statistical tests have not been applied to the results, though I would be happy to share my data with anyone with the skills and interested to do such tests.

In addition to empirical analysis, quantitative data and examples were produced through careful reading and rereading of posts on the Leaky Lounge, though the vast number of posts (450,000) means that still only a fraction of the posts on the Forum have been considered in this analysis. The Leaky Lounge is a public forum, and posts are as such not confidential. For those posts made in limited access forums permission to reproduce the material has been asked and granted. All names mentioned have been changed, with the exception of paperflowergirl (pfg), which is my screen name.

4. Results and Discussion

The following section presents the results of the quantitative and qualitative analysis, and offers explanations of why these indicators matters and analysis of the results.

4.1 Length

Post length was investigated in order to determine the relative frequency of the

other variables. As previously mentioned, all frequency indexes were produced by

dividing the raw frequency per twenty posts of a user by the overall linguistic output of

those twenty posts. The index for post length was calculated by dividing all scores output

by the number of words in the first twenty posts. The first index is thus one. The

subsequent scores are lower, indicating that overall input is slightly more than the

average for the first twenty post. However, aside from a slight initial drop, there is not

much change in the overall linguistic output over time. For this, and for all graphs for

which trend lines have been added, the trend line has been computed and inserted by the

Microsoft Excel automatic trend feature. As mentioned above, statistical tests have not

been performed, and as such the lines are only intended to be suggestive.
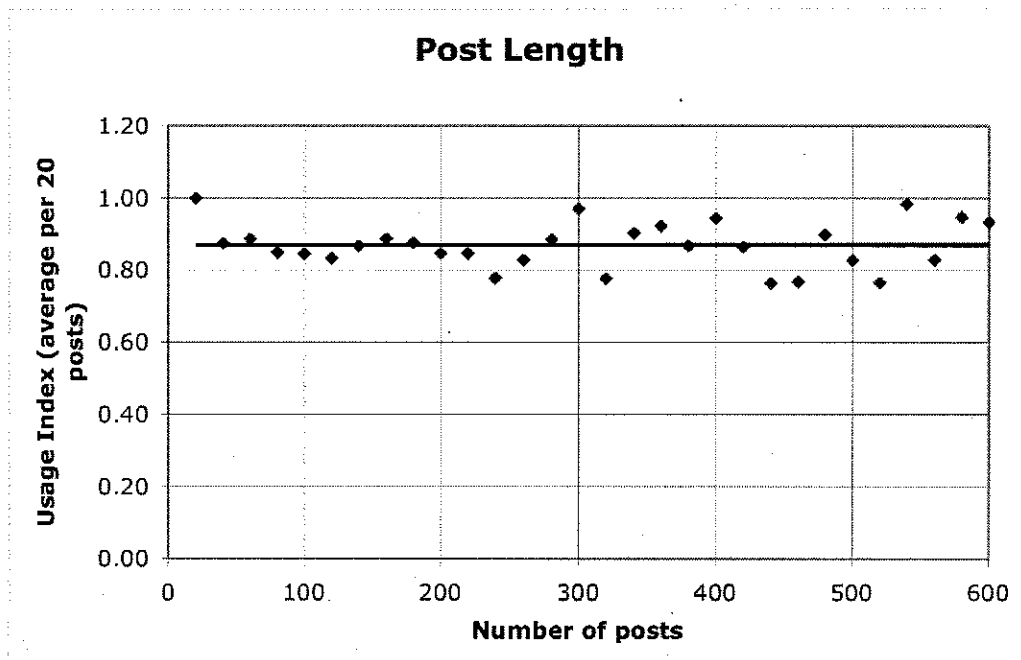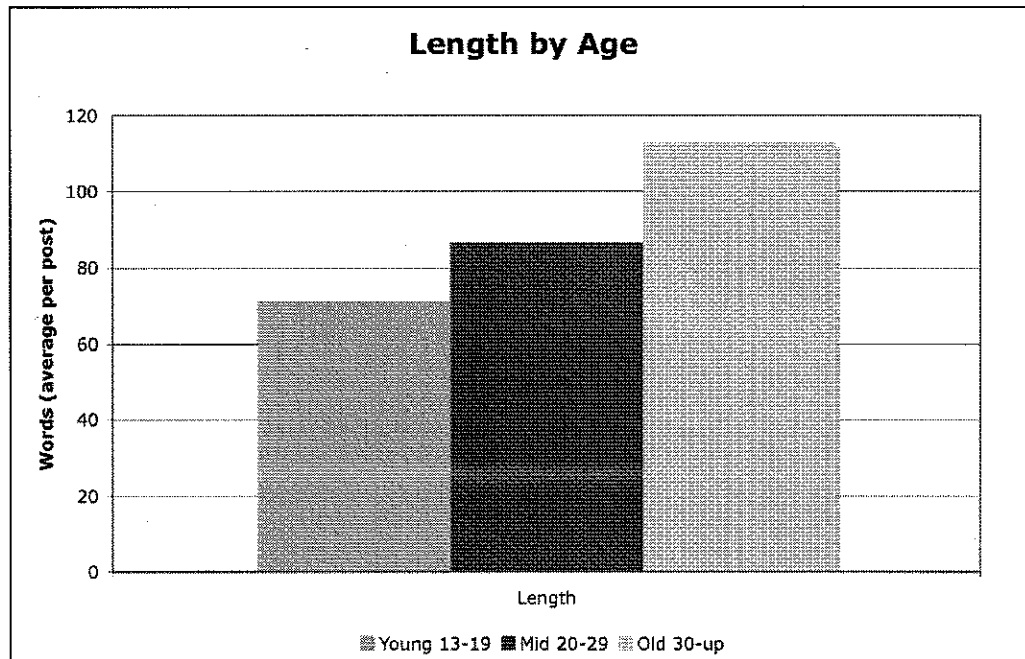
Figure 4: Post Length

Figure 5: Length by Age

Length by Age

Words (average per post)

Length

Young 13-19 Mid 20-29 Old 30-up

However, length is also a variable in its own right, which the results indicate depends primarily on age. Why this is the case is difficult to say, though it may reflect increased involvement in the discussions. More likely, however, it indicates a different kind of involvement in the discussions, given that younger members tend to post more frequently than older members, as seen in the graph below. The relationship between frequency and length was not clear from the sample of 15 members. However, a larger sample, including 70 of the top 100 posters (excluding those lacking profile information) shows, tentatively, that there is such a relationship. On these preliminary results it is difficult to tell whether older members are less active, or whether they rather 'concentrate' their time on the board into longer, rather than more numerous, posts.

Figure 6: Posting Frequency by Age

Posting Frequency by Age
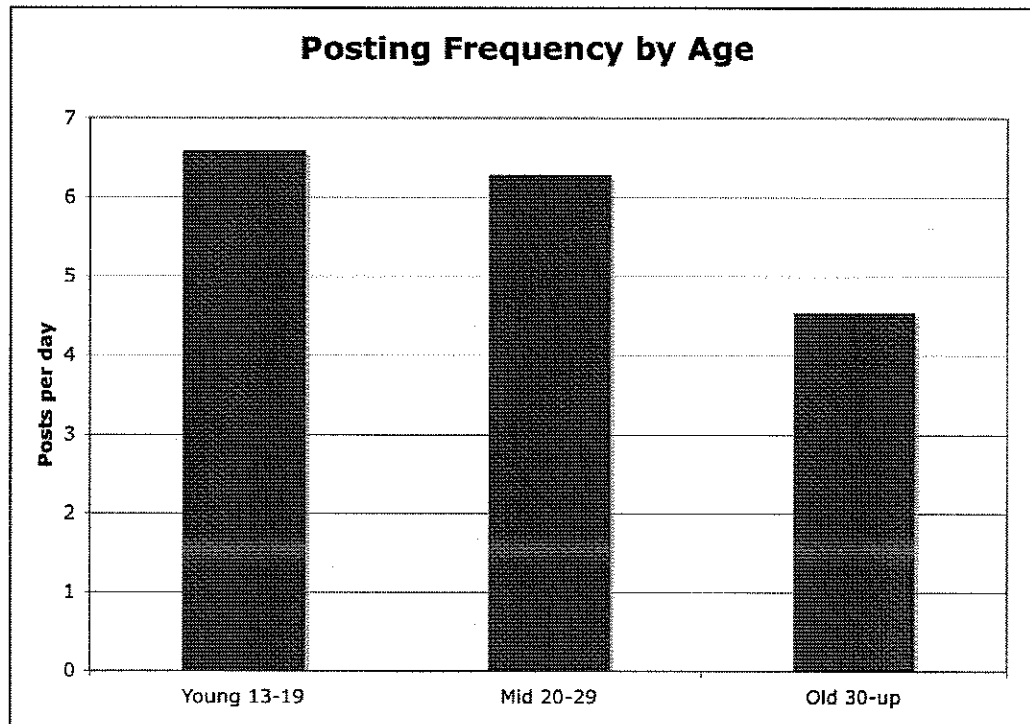
4.2 Linguistic Reinforcement of Community through Pronouns

The frequency of first person singular references (*I, me, myself,* etc) and first person plural (*we, our, let's*) over time should indicate whether there is a shift in how individual members conceive of themselves as members of a community. Cassell and Tversky hypothesized that if the Junior Summit '98 "was merely a gathering of individuals and not a true community, there would be no change in the participants' use of individual and collective pronouns," and found that there was indeed a significant change in pronoun use (2004:12-13). Thus one would expect *I*-references to decrease over time, and *we*-references to increase for the Leaky Lounge as well. The results, however, were not that straight-forward. This is explained by the fact that there are several differences between Cassel and Tversky's study and this one. Their study had the advantage of having a clear beginning and end, and the basis of analysis was thus time, not posts, which are unevenly-sized units of linguistic output. The other major difference

in the two online forums is that their forum had a set membership to which everyone posting had a claim. At the Leaky Lounge, however, membership is fluid, and posting does not automatically make one an insider or an *HP-sleuth*.[13]

The results seem to indicate that despite the friendliness of The Leaky Cauldron and the Leaky Lounge forum compared to a lot of online communication, some gradient distinctions of becoming in-group still apply. As mentioned earlier, these will be termed *newbie, novice,* and *wizard*. If an online community had a rising middle class, the *novices* would be it. They are members who have gotten past the technical and social obstacles of a community, and are eager to gain full member status, to be part of the 'in' crowd, so to speak. They are eager to distinguish themselves from the newbies, yet have not yet acquired an understanding of the unwritten rules and gained the respect necessary to pass for a wizard. In online message boards certain things, such as membership duration and post count are very visible indicators of status.

The Leaky Lounge has a policy of no harassment of new members, and the staff tries to be understanding when they make blunders with the rules. Unlike many forums, where membership status changes with one's post count, the titles awarded a certain 'class' of posters are not necessarily increasing in status. Member status will change from *Being Eaten by the Pea Soup*, to *Editor in Chief at the Daily Prophet*, but then be *Fighting off a Lovesick Shawl in the Secondhand Robe Shop*, and *Washing Windows at Weasley Wizarding Wheezes*. However, other forums not only use the member titles for thematic effect, but also to mark rank. For example, at the Chamber of Secrets Forum,

---

[13] This is a term coined by Galadriel Waters, author of *Ultimate Guide to the Mysteries of Harry Potter*, and a range of other books aiming to decipher the clues Rowling drops in the books and predict future books. An *HP-sleuth* is someone who knows the books very well, but can also create analyses and especially theories and (wild but somehow believable) speculations.

members start out as first years, and progress to second years, third years, and so forth. Once past 300 posts, they may be invited by the staff to join Hogsmeade, an exclusive forum for more experienced members.

The results, especially those of pronoun-references, thus reflect two trends at the same time: first, the desire to be perceived as an *HP-sleuth*, and second, increased linguistic reinforcement of community.

4.2.1 First person singular references

First person references over time, such as *I, me, my, myself, I'll, I've, I'm* and so forth, show a trend of increase. The best explanation for this unexpected result is that the increase in first person references correlates to the increased conversational and interactive style of the members. First person references might be taken to indicate increased self-absorption, the opposite of community-reinforcement. As already mentioned, Cassell and Tversky recorded a drop in first person singular references (2004;13). However, there is a basic difference between the Junior Summit '98 and the Leaky Lounge: the former was a task-oriented forum, whereas the second has more a social orientation. Leaky Lounge members do not have to accomplish anything. Thus, as familiarity increases, so does the individual's willingness to put him or herself 'out there,' in the form of expressions about personal information such as feelings, reactions, RL events, and so forth. This interpretation is corroborated by the results for *we*-reference, *you*-reference, and a range of interaction-markers discussed later.

Figure 7: First Person Singular References

First Person Singular References
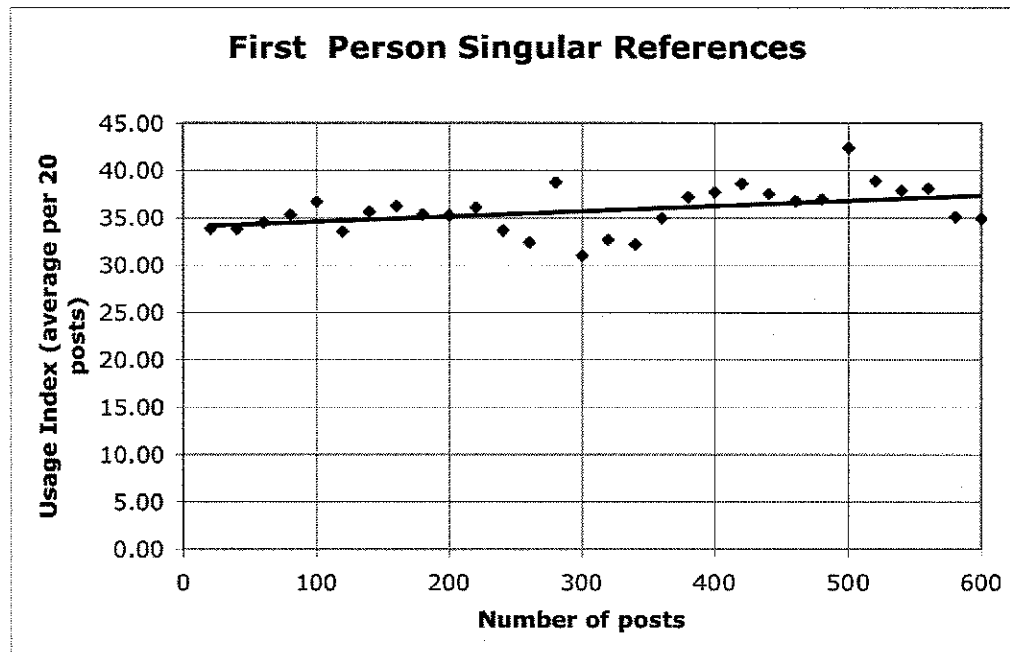
4.2.2 First person plural references

Similar to Cassell and Tversky, Wick has also examined the use of *we* to constitute community. She argues that the members of Pednet used *we* constitutively to determine their identity as a group, distinguishing themselves as political walkers as opposed to recreational walkers, but accepting bicyclists as part of the group (Wick 1997;62). Similar usage can be found on the Leaky Lounge, where members use *we* and related first person plural references to refer to various sets of overlapping groups: the participants in a particular discussion, members of the Leaky Lounge, or Harry Potter fans in general. In most cases, this distinction does not need to be clarified, because these groups are, as said, overlapping, and the referents are generally clear from context.

1. We cannot decide whether Snape is good or evil. This becomes very clear, because despite all the wonderful comments here, we are still in the dark.

In the above example, *we* refers to the participants in the discussion, but it is not inconsistent with fandom in general. The implied referent of *we* below is HP fandom, or HP readers.

2. Since there's more to be revealed about so many others, we may very well not know Peter's whole story either?

However, sometimes members do clarify what they mean. This shows that at least some of the members are conscious of who they consider part of their community. It is interesting to note in the example below that the fictional characters populating Rowling's universe (the Wizarding world) are explicitly made part of the group. The explanation for this lies in the fantastical aspects of the HP online community, which allow the referential borders of what is real or not to be blurred. Further discussion of this will be continued in section 4.8.

3. I, too, kinda think that Kreacher may have been the "traitor" this time round in the series. He's already caused a death which we all took pretty hard ("we" being some people in the wizarding world and the readers of HP).

The results show an increase in first person plural references over the span of 600 posts.[14] The data is quite scattered, but illustrate a few things: First, the overall increase in the span of 600 posts confirms the increased perception and reinforcement of community by the members. Second, the relatively low level of references during the first twenty posts also indicate that new members are quite cautious at first to use the collective, but already in posts 21-60 they feel 'entitled' to use such references more.

---

[14] A short note should be made here to explain that these results are for aggregate usages of *we*-related statements. However, these results were not computed manually, and as such these results also include first person plural statements not referring to HP fandom, but to other non-inclusive groups, such as the member's family.

However, more constant integration of first person plural references takes considerably more time.

Figure 8: first person plural references

First Person Plural References
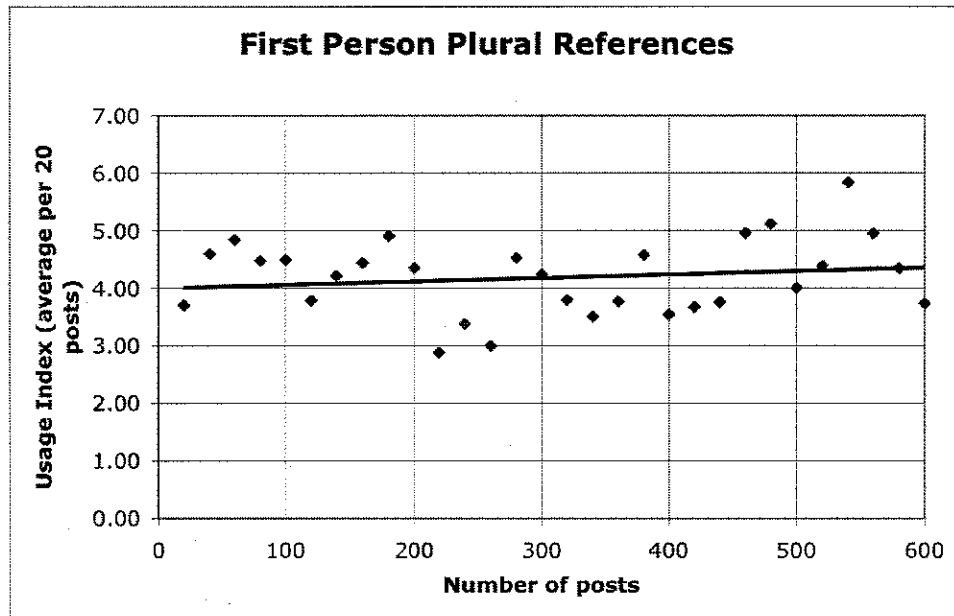
4.2.3 *Pronouns and the prestige of being an HP-Sleuth*

However, there is also another interesting feature of the change in *I-* and *we-* references over time, which is a lower than expected level in such references between 200-400, with a low centered around 300 posts. This dip, also present for second person references (see section 4.3.1, Figure 11), is most probably motivated by the desire to prove one's worth as an HP-sleuth. There is of course no obstacle to being conversational and sleuthing at the same time, but being less conversational sends a certain message: the flip-side of being less personal and less conversational in one's language is that one appears more formal and/or academic. To prove their skills as knowledgeable HP-analyst, members are thus reverting to more formal language.

Though not always unequivocal, this trend can be seen in all the graphs of pronoun

references, and is also apparent on the aggregate graph, below.

Figure 9: Total pronoun references

Total Pronoun References
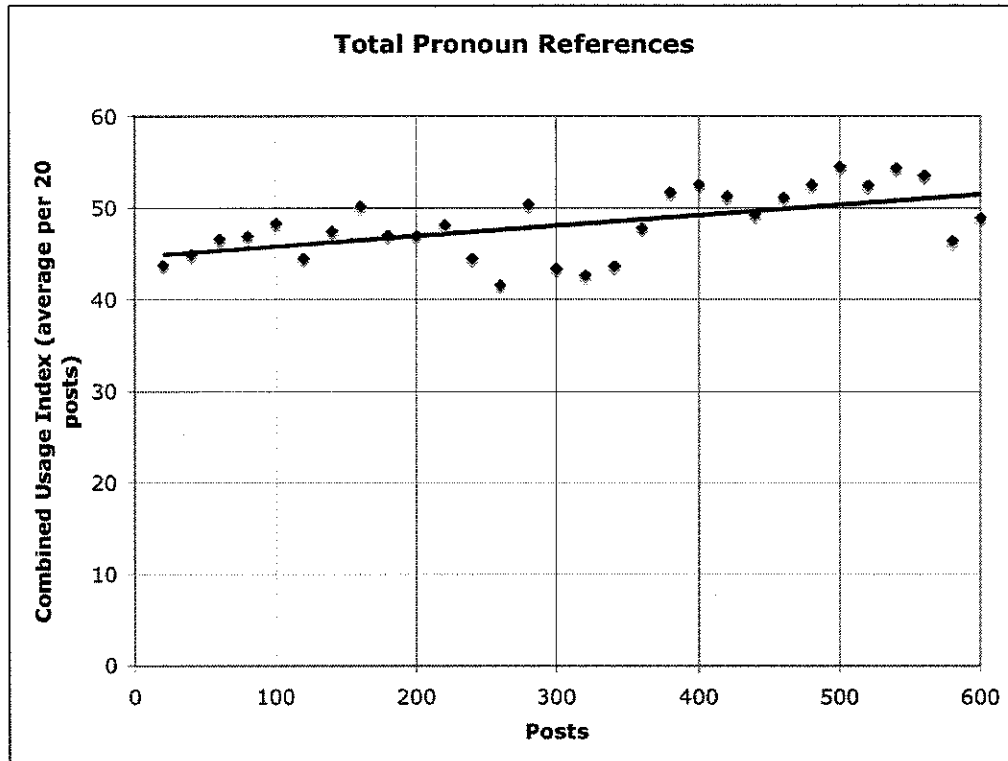
*4.2.4 Pronoun references and age*

When comparing pronoun use to other demographic indicators, such as age,

gender, and frequency, one trend becomes clear: There is a difference in usage between

younger and older members, with older members using less *I*-references and more *we*-

references, and younger members doing the opposite.

Figure 10: Pronoun Reference by Age

Pronoun Reference by Age
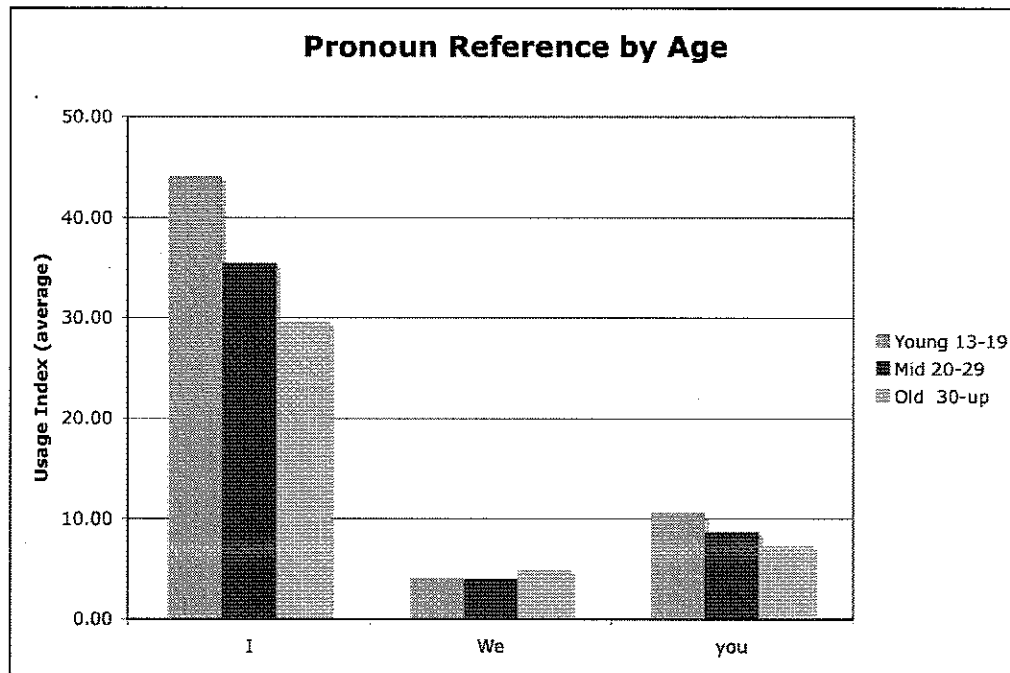
This trend is consistent with the findings by Pennebaker and Stone (2003), who have studied changes in speech related to age. In their study of written and spoken (later transcribed) statements by over 3000 research participants, they found a considerably increase of the first person plural due to age (Pennebaker and Stone 2003:296). They also found a sharp decrease in self-references with increased age (Pennebaker and Stone 2003:297). The decreased use of second person references due to age, in the context of a message board, might indicate a less conversational style among older members. However, Pennebaker and Stone, who found lower levels of references to other people with increased age, believe that this may "reflect a greater independence and focus toward the external world" (Pennebaker and Stone 2003:299).

4.3 Interaction

As mentioned earlier, CMC communication on message boards need not necessarily be interactive. Though it is frowned upon, members can, and sometimes do,

post their own ideas and thoughts on a particular topic without reading a word of what anyone else has said. In a RL conversation, it is not as easy to ignore another person's opinion. Even if person A chooses to ignore what person B says, B can be confident that A (in normal circumstances) did at least hear what they said. In many forms of CMC, perhaps particularly message boards, not 'hearing' is an option. Being heard is important, and some go so far as to claim that "unless one's contributions are acknowledged by others, one does not actually 'exist' online" (Wick 997:6). In real life, too, is the act of being ignored hurtful. However, in CMC, expressed recognition is perhaps more important, given the absence of non-verbal affirmations. While this 'lack of existence' may hold true to a larger degree for smaller groups, such as mailing-lists, some form of validation is important also for participants in message boards, in the absence of eye-contact, nods, and *uhum*-statements. To respond without first reading is unpopular, and clearly subversive to a true community. Since the basis for online communities is overwhelmingly verbal in nature, interaction between members is a necessary component. It is also one that can be expected to increase over time, as the community becomes stronger and members adapt to the requirements of CMC.

Several markers of community interaction were investigated. Whereas manual coding of posts allows an open approach as to what phenomenon are present in the data, this study has relied upon the findings produced by Cassell and Tversky (2004), who also looked at instances of agreement, disagreement, greetings, and Wh-questions. To these have been added second person pronominal references, questions in general, and quotes.

4.3.1 Interaction through second person pronouns

The previous section explored the use of *we* and *I* by Leaky Lounge members. However, the sharpest increase in pronoun use can be seen in second person references. Such *you*-statements were not measured by Cassell and Tversky (2004), nor by Wick (1997). However, they provide a simple way of ascertaining to what degree members engage with each other.

Second person references can fulfill two separate purposes in the conversation: the generic *you*, and the specific *you*. Examples of such usage can be seen below. The first sentence makes use of the generic *you*. The second and third exemplify interactive behavior between members. The fourth sentence, whose type is considerably less common, is noteworthy because it is directly addressing one of Rowling's fictional characters, and serves to reinforce the fantastical aspect of the Leaky Lounge.

4. The worst part about broken toes is there is nothing you can do about them...
5. Also Belissa, your comments that you have posted have been thoughtful, and heartfelt.
6. You guys have all posted excellent theories as to what that gleam might be.
7. Also Happy birthday Ginny!!! [birthday-cake emoticon] I hope you're celebrating with Harry :)

Figure 11: Second Person References

Second person references
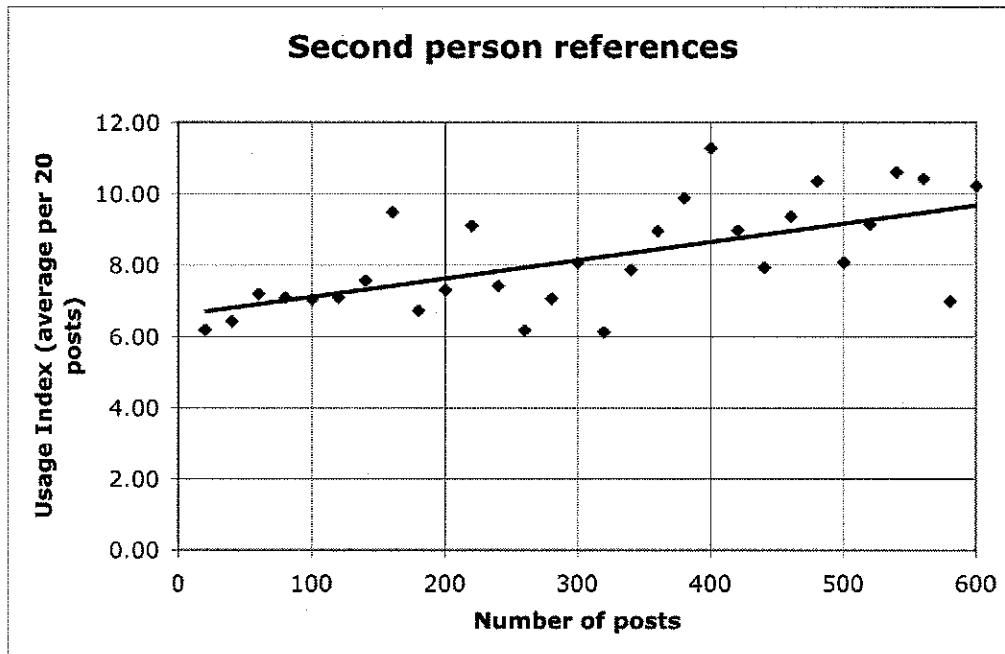
The increase in second person references thus indicates both an increase in conversational style, and an increase in interpersonal interaction. As discussed above, there is a dip in usage around 300 posts, marking a more 'professional' approach by members seeking to establish themselves as HP-sleuths. However, the overall appeal of *you* as a means of reinforcing and connecting the community makes this trend less visible than for *I* and *we*.

4.3.2 Interaction through Agreement and Disagreement

Through searching the data for the most common expressions of agreement and disagreement, I sought to uncover what variables influence the use of these interaction-markers. Agreement and disagreement can play very complex roles in fostering community, and is often combined with the use of quotes. The Leaky Lounge, like most message boards, has a feature that allows a member to insert a box containing all or part of another member's post, and may be repeated, producing a box-inside the box effect.
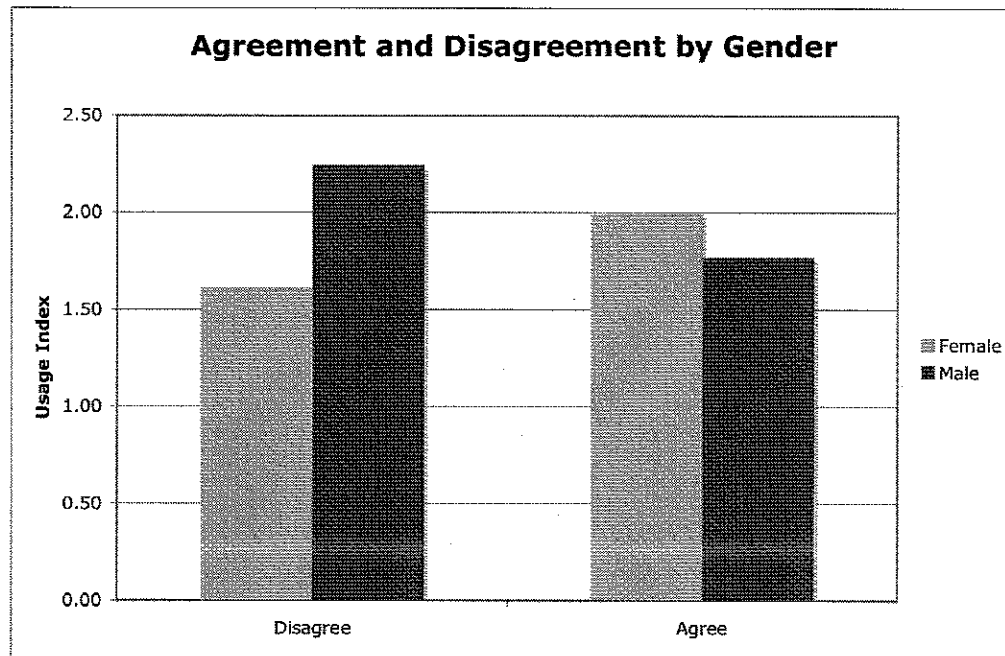
8. I agree with much of what you say, McGonagall's Cat. :) Except that I don't believe Fudge's explanation on why he brought the Dementor with him. See below.

Figure 12: Agreement and Disagreement

Agreement and Disagreement
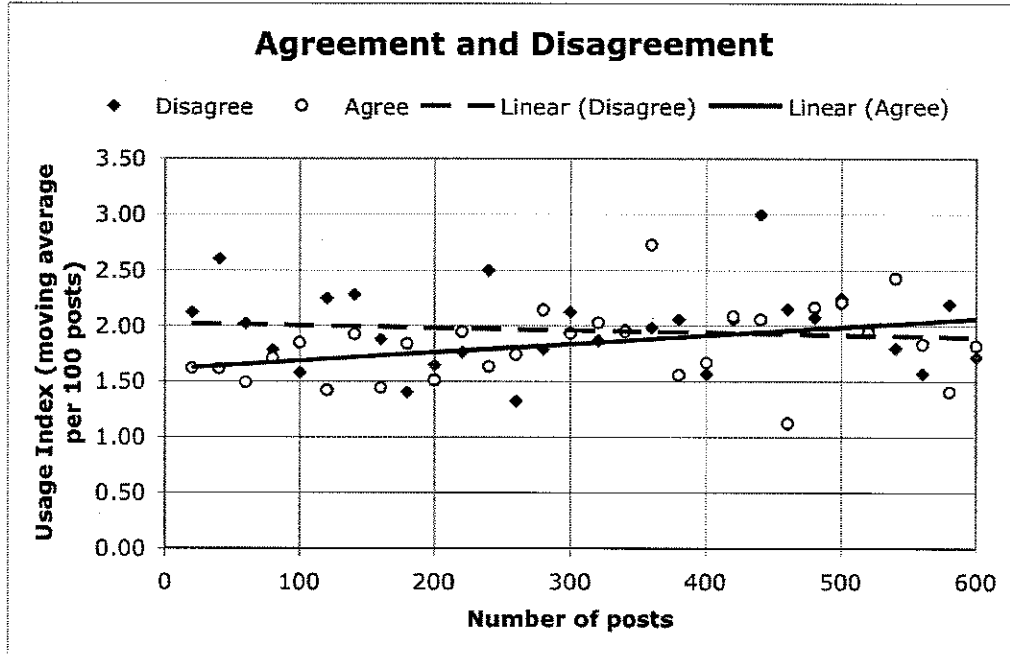
It was projected that instances of agreement would increase, and that disagreement would decrease. The results show that the overall level of agreements increased over time, validating the hypothesis that as members become more integrated and gain seniority in the speech community, they also reinforce the community through increased supportive interaction with other members. There was a very slight decrease in disagreement. This may be caused by the fact that disagreement, while negative in content, also builds community because it is a form of interaction.

Figure 13: Agreement and Disagreement by Gender

Agreement and Disagreement by Gender

A chart with Usage Index on the y-axis (0.00 to 2.50) and two categories on the x-axis: Disagree and Agree. Legend: Female, Male.

A rather striking, though perhaps not surprising, result can be seen when calculating agreement and disagreement by gender. The results show that females are 'friendlier' in their communication. Wardhaugh notes that there are often different politeness requirements placed on males and females, and that women's language tends to be more conciliatory (2002:319). While traditional forms of reassurance, such as *mhmm* interjections, are not possible online, explicit statements of agreement are. Wardhagh also points out that males have a more competitive language (2002:323), which in turn explains the higher disagreement score for males.

4.3.3 Interaction through Greetings

Greetings are another measure of interaction between members that also indicates how eager a member is to introduce him or herself, but also how willing the member is to welcome new member once they are past their own introductory stage. However, the presence of greetings in a person's vocabulary seems to be to a large extent a function of

personal preference, since the variation of greetings ranged from 83 to 7 per 1000 posts.

Also, the relatively low frequency of greetings make the results less clearly indicative.

Figure 14: Greetings
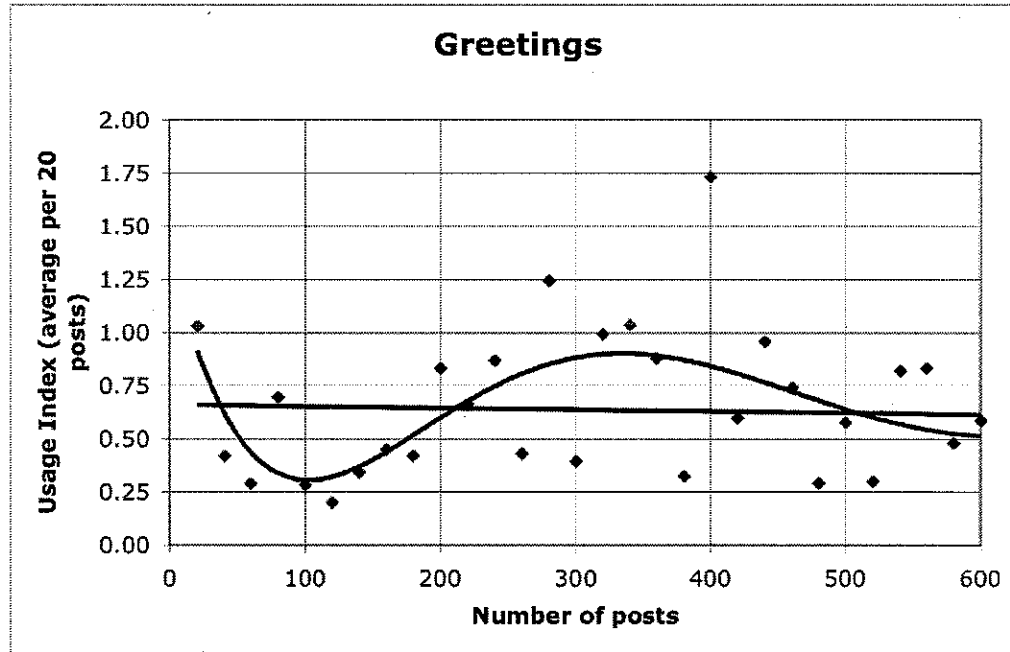
The greeting results are quite scattered, but a linear trend line shows that the

scores stay quite constant in the long run. The reasons for this are not clear, but the very

low frequency of data may account for the fact that no long term trend is clearly

observable. However, a more flexible, polynomial trend line captures the rapid and quite

uniform decrease in greetings after the first 20 posts, which does not pick up again until

about 200 posts. The initially high level of greetings is explained by members wishing to

introduce themselves in the first few posts, as in the following example.

9. Hi, I'm Linga, and I'm really tired at the moment so I shouldn't post right now. :D

However, many members choose to simply join the conversation without any

form of introduction, even in their first few posts. Later, members may (very

47

occasionally) say *hello* when joining an established conversation, but mostly greetings

are reserved for responding to new members, or to notify others that one will be away for

a while.

The following drop in greetings has two related explanations: first, members are

not yet 'senior' enough to welcome other members, and second, may indeed not wish to

be too friendly with new members because they are trying to distance themselves from

the newbie category.

Results for greetings show unusually clear results for gender, with females

greeting more than 2.5 times as much as males. As with agreement, this is also to be

expected given what has been observed about gender variation in language.

4.3.4 Interaction through Questions

Questions were measured through the proxy of questions marks, and Wh-words

preceding question marks. There are thus two potential errors of measurements: first, the

use of excessive punctuation marks to mark one question (as opposed to several), which

is not uncommon in CMC.

10. But--if there was already a part of his soul hidden in the diary, why did he need
Ginny's soul to step out of the diary and become more than a memory???

Second, the Wh-question score is likely to be higher than actual Wh-questions,

because my knowledge of coding was not refined enough to exclude results such as the

sentence below, where *where* precedes a question mark, but the pair do not constitute a

Wh-question.

11. That's where Harry and the Weasley's live. Think the Gryffindor item is in the right location?

The results for this marker were not clear, showing a slight decrease in Wh-questions over time, and a slight increase in the use of questions marks. It is not clear why overall questions would increase while Wh-questions decrease at the same time, since they fill very similar functions. Most probably more refined methods of inquiry will be required to measure these variables accurately.

4.3.5 Interaction through Quoting

One measure of interaction is the degree to which members quote each other. Those who have published works can understand what it means to discover that someone else has cited their book or article. The quote feature, whereby a member can excerpt part of the message of a previous poster and include it in their own post, is a similar way of giving credit. It is also useful in keeping the important or insightful points part of the discussion. Lastly, they show that members are indeed interacting, in terms of ideas, but also socially and for humor.

12. On topic post with quoting feature (post by Henry)
    >Quote (pfg @ Mar. 21 2005,10:44 am)
    >[pfg] A person incapable of good cannot be justly punished for doing evil.

    [Henry]O.k.. But, as Galen wrote (quoting from memory): "We do kill snakes that did bite although we don't assume that they have any freedom to bite or not to bite. We kill them in order to prevent that they will bite again." And I'm not sure that we are doing the snake a favour in regarding (and treating) it like this.

13. Humorous post with quoting feature (post by Henry)[15]

---

[15] The humor in this conversation stems from the fact that MysteryMan's point about Draco narrowly escaping Muggles by helicopter is false: those familiar with the books know that Draco is lying when making this statement. However, Henry chooses to play along with MysteryMan's false information. *Mudblood* is a derogatory term for witches and wizards with Muggle (non-

>Quote (MysteryMan @ April 12 2005,11:31 am)

>>Quote (Henry @ April 12 2005,05:05)
>>[Henry] Draco Malfoy hates "mudbloods", but probably not because of
having been enraged by muggles, but because he's been brought up like this.

>[MysteryMan] What about all those narrow escapes from Muggles by Helicopter
;) (PS/SS)

[Henry] Excellent point! [laughing emoticon] Hadn't thought about it. Sorry!
[emoticon banging head against the wall]

This second exchange shows a post by Henry quoting a post by MysteryMan, in
which MysteryMan in turn quotes Henry. It thus gives a history of this particular
conversation, since not all posts in the discussion (thread), nor all information in a
member's post will pertain to this particular point.

Unfortunately, scores for the use of the quote feature were not mined, because of
the intricacies of detecting the coding features. However, observational evidence seems
to indicate at least a shift in usage, from merely quoting another person and adding *good
point*, to the more complex kinds of interaction seen above, where members further refine
the ideas of the previous posters, or use the quote feature to neatly execute a joke.

Overall, then we see that there are several ways in which members of the speech
community can signal acknowledgement and to one another. The qualitatively measured
markers tend to increase as members gain in seniority, though some, such as second
person pronouns and greetings, may experience a period of decrease as novices are trying
to 'prove' themselves.

---

magic) parents. *PS/SS* is a citation of the book, *(Philosopher's Stone/Sorcerer's Stone)*, in which
Draco says this.

4.4 Friendliness

Two measures of friendliness, or politeness in conversation, were investigated: apologies and personal qualifiers. CMC is more prone to conflict and misunderstanding (which sometimes lead to flaming and flame wars) because there is a lack of visual and tonal reassurance of friendly intentions. However, there are also verbal cues that can indicate friendly intentions, and make a speaker come off as less abrasive. One way of making other members feel welcome in the discussion are the already mentioned interaction variables; by showing that you have paid attention and care what other people have said. The two ways investigated in this section were apologies and personal qualifiers. Apologies are conciliatory measures that show concern for one's behavior relative to the rest of the community. Personal qualifiers, such as *in my opinion*, may serve to 'soften' the edge of one's argument. They thus constitute friendly language.

Unfortunately, the relative frequency of apologies captured by the search was too low (averaging less than once per 60 posts for most) to produce any clear results given the sample size. An inherent difficulty for both these inquiries is that the context that would otherwise clarify the 'tone' of the apology and qualifier cannot be considered. Statements such as *"excuse me, but"* have polite wording masking a negative comment.

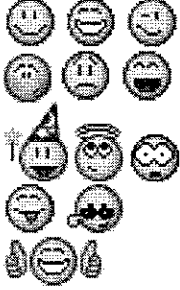## 4.5 CMC Adaptability

Measures of the use of particular CMC features are slightly different from the above variables, because most, if not all of these, are not only indicative of the linguistic behavior of Leaky Lounge members, but also of their comfort with the computer medium. Unlike variables such as pronouns, greetings, agreement and apologies, all of

51

which occur in spoken and traditional written language, the CMC features do not have a non-computer equivalent. They thus have to be learned.

In learning to use the full range of options on a message board, there are several features to be considered that differ from other kinds of CMC and traditional writing. They range from using emoticons, stars, and abbreviations to puncutuation and the increased graphical flexibility of message boards, including capitalization, italics, colors, and spacing. These options are described in more detail in the CMC table below.

Table 5: Description and examples of CMC features

Description	Examples
Emoticons or smiley faces In the early days of computer communication, such as Internet Relay Chat (IRC) communication, emoticons were combinations of standard keyboard strokes that combined to make faces with different expressions when seen from the side. They are still used as text in e.g. e-mail, but many types of CMC facilitating software, such as Instant Messenger (one-on-one synchronous chat) and message boards now replace these with more visual emoticons that use colors and sometimes also movement.	Smile :-) and :) Grin :D Wink ;-) and ;) Tease :-p Grim :-| Sad :-(
Netspeak Netspeak is a large category of words, expressions, and abbreviations that convey information quickly. Netspeak may arguably include all other linguistic forms particular to CMC, depending on one's definition. Some of these act similarly to emoticons, such as LOL (laugh out loud) and thus convey mood. Other expressions, such as IMO (in my opinion), OMG (O My God) , or BRB (be right back) convey common information.	*Lol*: Laugh out loud *Rotfl*: rolling on the floor laughing *IMHO*: in my humble opinion
Onomatopoeia Several forms of spoken language have not traditionally had written forms, but for the purpose, first, of comic books, and later of CMC, non-lexical expressions are being written down.	Hehe, Mwuahaha, oh, huh? Mwah! Hmmm… Wheee! Whoa! Yum! Mmmm, Nihihi, hahaha, Ha! Eh.
Punctuation Punctuation is used similarly, and sometimes in conjunction with, onomatopoeia and text formatting, as a proxy for spoken tone and emphasis. For example, *Hi…* and *Hi!!!!* convey two very different levels of excitement (which pitch and volume would usually provide in real	Hi……. Hi!!!!!!?! I'm not so sure.… No!!! *Nooo!!* Yes… Yes!! YE$$! Yes? huh? WHAT????? Am I right or am I right??!?!?!

life) on part of the speaker.	
Starred action Users of CMC get around the limits of the screen, on which no body language is visible, by using stars (*) or sometimes colons (:) or angle brackets (<>) to highlight information about what they are doing. Such statements are always written in the third person, with the author being the implied subject. Actions do not have to be reasonable, nor realistic, or related to real life objects. Virtually served coffee and Krapfens make just as good references as RL monitors and grumbles.	*grumbles*, *picks up a couple of krapfens and runs to catch the bus*, *snuggles up next to MysteryMan and falls asleep on the couch*, *drools in her coffee*, *glares at the monitor in disbelief*
Text formatting and graphical flexibility Because early CMC media did not include features such as bold, italics, and underlining, these tended to be expressed through other symbols. These serve to add emphasis. For example, using capital letters is the equivalent of yelling, and excessive use is perceived as equally rude as RL yelling. Small letters may be perceived as whispering or mumbling. Bold, italic, and underlined words draw their meaning from traditional written language, where they are used for purposes of emphasis.	*bold* _italic_ **bold**, *italic*, underlined, strikethrough, YELLING, colored letters, Large, small, Font-variation
Quote feature This is part of the board coding features, and lets members include all or parts of other member's posts in their own posts.	>Quote (Persephone @ June 15 2005,5:34 pm) >Henry, we'll agree to disagree. Yes. :) No harm in that.

As members learn to master these new forms of expressions, their ability to do many things increase. Emoticons, for example, are very important in reassuring others of one's friendly intentions, especially sarcasm.

14. Werewolf! will you quit followin' me around! :p [teasing emoticon]

In the above example the inclusion of the emoticon is very important. Without it, the speaker would easily be perceived as rude. The teasing emoticon (sticking its tongue out) makes it explicit that the sentence is in jest. The member may indeed appreciate having Werewolf posting in the same conversations as her. Below, the same effect is produced by the emoticon, acting to reassure the readers that the intentions are friendly.

The triple period marks serve to illustrate this member's hesitation in finding a way to describe the other participants in the conversation, acting as a stand-in for a verbal pause.

15. Oh! . . .and I thought I was crazy! You guys are. . . uh . . .
    *backs away to go join slightly less ~~psycho~~ nutty group . . . *
    :D

16.
    Post 1: **Herb**:   Welcome to the new mods! And I hope all you sick folk ~~stay away from me~~ get better soon.
    Post 2: **Sparrow**:   *sneezes on Herb*
        wipes beak on Herb's sleeve
        I love you Herb!!!
        Why are we all of us ill?? And at the same time!!!!
    Post 3: **Sim**:   *sniff*
        It's YOUR fault---you're spreading Avian Flu!!!!! Oh wait, that's just to me. No idea then :)

The crossing out of text by Herb in example (16)does not have a direct equivalent in RL communication, but the closest would be slips of the tongue. However, since forum posts are instantly retractable, real slips of the tongue are impossible to mimic. Sparrow, having a bird identity, uses the fact that she is sick in real life to get back at Herb, but modifies it to fit her online persona. Sim catches the humor of the situation by blaming Sparrow for spreading the avian flu, which featured much in the news in the weeks before this conversation. She also halfway retracts her statement and adds a smiley-face to show that she has friendly intentions. This is therefore an example of members who are skilled in written CMC, and use a range of net-speak features to discuss sickness in a humorous way, mixing RL information with virtual reality and online communication features.

Keeping in mind some of these opportunities created by 'rich' CMC interaction, especially for maintaining friendliness in social interaction, and increasing the range of

'safe' expressions to include humor and irony, it is not surprising that the these features increased over time. Not all features discussed in the CMC table were measured. I chose to focus upon emoticons, Netspeak, text formatting, and starred expressions. Two of these variable showed a clear increase overall, and two showed a small increase. However, the time frame during which this happened is not uniform, showing that different features have different learning curves.

4.5.1 Emoticons

As discussed above, emoticons are important tools of communication because they serve to clarify intentions, especially those of friendliness. The data shows an overall increase in emoticons over the 600 posts covered by this study. However, the major increase in usage of emoticons happens in the first 80 posts. For the first 20 posts, members use about half the amount of emoticons they later employ. This rapid change during the first 80 posts is probably the result of an emoticon 'learning curve,' during which members learn through observation, and occasionally through facing the consequences of not making one's friendly intentions clear enough. An example of such a case can be seen below, where a member apologizes for having been misunderstood. This time when she reiterates her point, she includes a smiling emoticon.

17. Um, I really didn't mean to offend people by saying I dislike The Rasmus, but this thread is for naming dislikes... :)

Figure 15: Emoticons
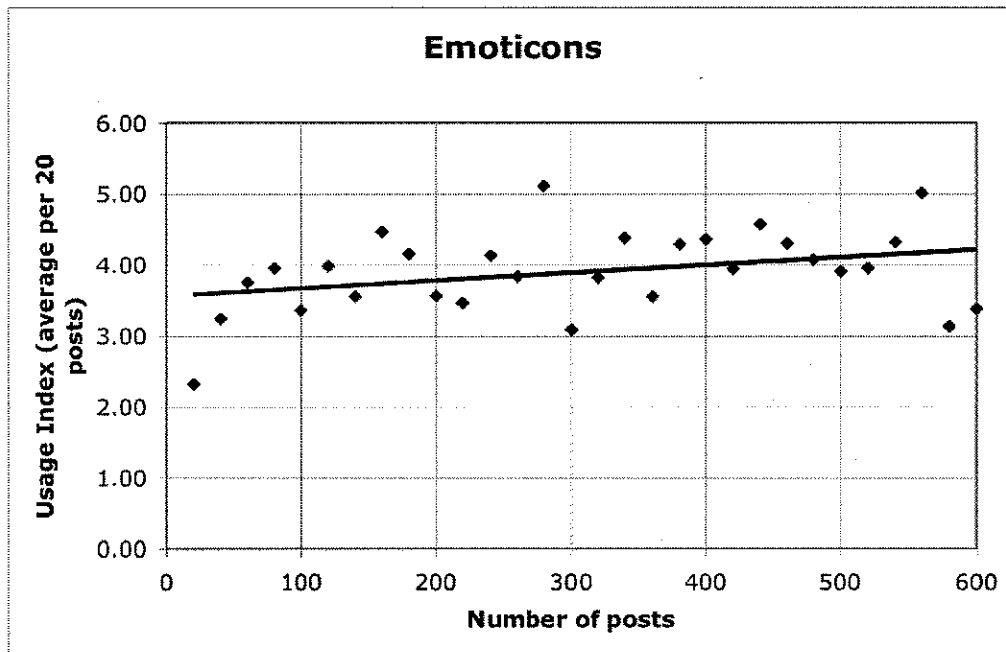
Emoticons

4.5.2 Netspeak

The use of Netspeak statistics sought to capture member's acquisition of common CMC

terms and abbreviations. Netspeak is not as uniformly community-building as emoticons,

since a large portion of Netspeak terms are merely abbreviations for common phrases,

such as *by the way*. However, some Netspeak terms, especially those related to humor,

serve a similar function as emoticons: they signal intent. Below the accusation that the

previous poster was rambling is softened by an added *lol* for *laughs out loud*.

18. Hahaha! This sounds like something else... are you trying to make a point here? Lol!

The increase in Netspeak was quite small. This is probably because Netspeak,

while spreading some between members of the board, still remains a matter of personal

preference. Average scores per member range from 0.4 to 5.7, which is quite a wide

range. The use of Netspeak is also a matter of personal risk, because it is formally not

allowed on the Leaky Lounge. However, this is a rule that is only mildly observed (even

by moderators) and enforced only in the more extreme cases. The unofficially recognized range of usage appears to be limited to laughing statements, such as *LOL* (laugh out loud) and *ROTFL* (rolling on the floor laughing), and a few personal qualifiers, such as *IMO* (in my opinion) and *IMHO* (in my humble opinion). Members wishing to appear responsible will therefore minimize any extravagant use of Netspeak. A closer examination of what forms of Netspeak tend to be tolerated in practice, and a re-calculation of usage indexes based on these might produce stronger results.

4.5.3 Text Formatting

The frequency of text formatting, such as bold, italics, and colored text was also measured. Such features, while available in traditional written formats, are an integral part of the how a member expresses emphasis and draw attention to certain parts of their posts. The results indicate an overall increase. The increased use of formatting thus signals increased ability to use the CMC medium effectively. Increased use of formatting may also be seen as an indicator of increased willingness to take the time to make messages clear and expressive, that is, a greater commitment to the readers of the message. Further discussion of text formatting, including its status as both a CMC and traditional writing tool, can be found under the discussion of CMC features by age and gender in section 4.5.5 below.

Figure 16: Text Formatting

Text Formatting
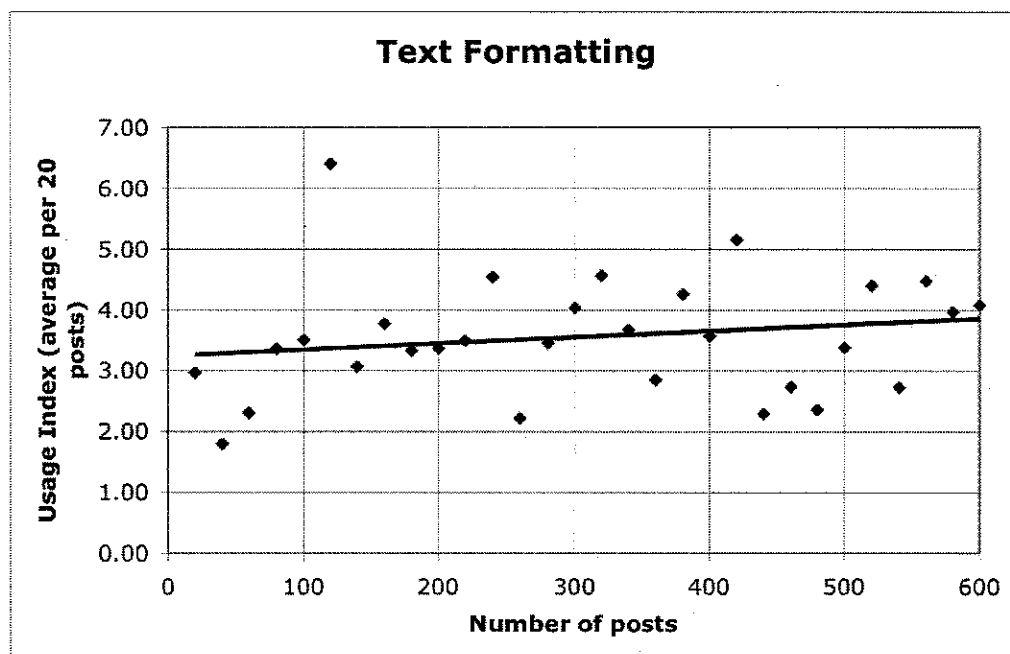
4.5.4 Starred Action Expressions

Placing action within two stars or other form of indicator is a kind of communication with no parallel in traditional spoken or written language. While some actions are indeed reflective of real life prosody (such as facial expressions and gestures), these expressions differ considerably because they do not in fact have to be real, or even possible. Nevertheless, they can be very expressive of a range of feelings, including those who are not covered by emoticons. As seen in the Avian Flu example above, action expressions do not have to be expressive of emotion.

19. I still can't believe he didn't go insane after being tortured by Voldemort in GOF (*shudders just to think of it*).

20. Please! ...*blush* This is a family site! :D

The data indicates an overall increase in the use of starred expressions. Like the other variables, this increase shows increased adaptation to the CMC medium. However, the increase is not as sharp as that of emoticons (where the first 80 posts were the most

crucial), indicating that members take longer to adopt this feature into their vocabulary. This is probably because the usage of such stars is more complex. Emoticons, at least the simple, non-moving ones, generally act as proxies for facial expressions. However, the range of what stars can model and convey is much larger. Though the basic concept of an action being performed, not only facial expressions (*grins*) can be conveyed, but also onomatopoeia (*coughmisogynisticcough*), thoughts (*shudders just to think of it*), feelings (*sob*), RL actions (*goes in search of chocolate*), virtual interaction (*gives Sparrow a cup of hot chocolate*), straight to the fantastical (*conjures up a Butterbeer fountain*). Thus the range and complexity of their usage require time to adopt.

Figure 17: Starred Action Expressions
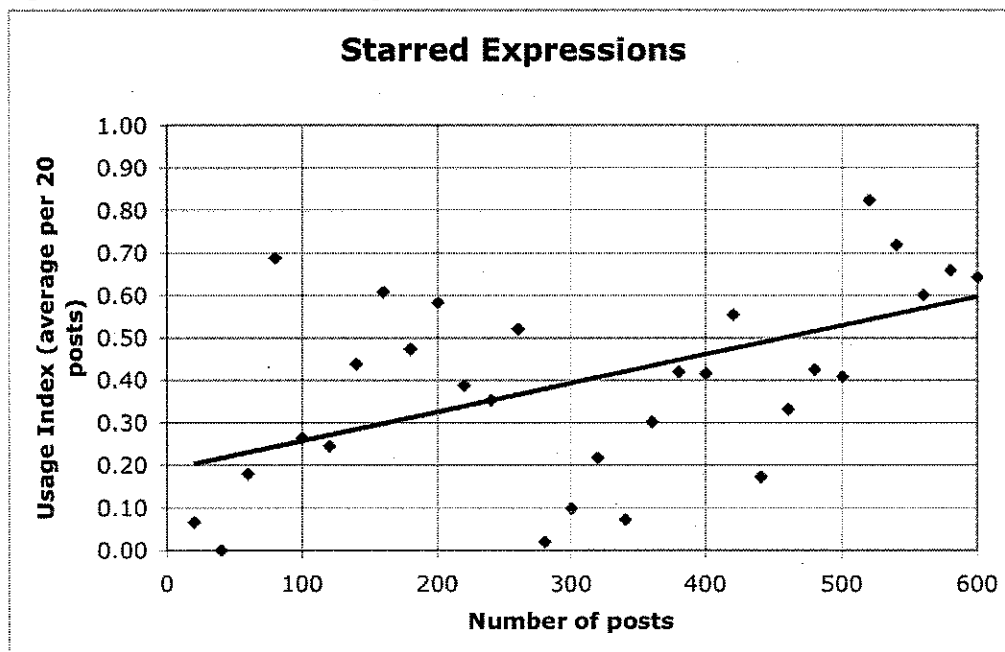
As mentioned in the CMC table, stars also serve the function of highlighting words. A closer (manual) look at the data revealed a gradual shift: early posts tend to use stars to give emphasis, and later posts use them more for action statements. This shift

away from emphasis usage can be explained in part by the fact that the board has bold,

underline, and italic functions that can be used instead.

21. ...Slytherin house is the only house concerned with blood and therefore the only one
    likely to make sure that there is an *heir* born...

4.5.5 Age, gender, and CMC features

The influence of gender and age on CMC usage was very strong. Almost without

exception, females scored higher than males, and young speakers scored higher than old

speakers.

Figure 18: CMC features by Gender
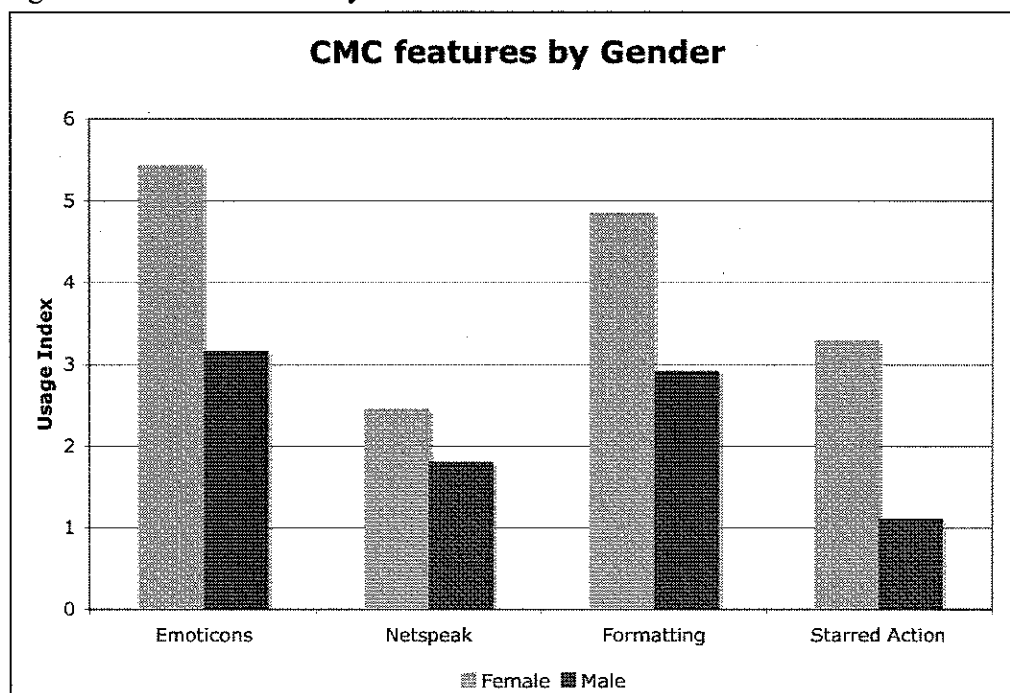
Figure 19: CMC features by Age

CMC features by Age

Usage Index

Emoticons   Netspeak   Formatting   Starred Action

■Young 13-19 ■Mid 20-29 ■Old 30-up
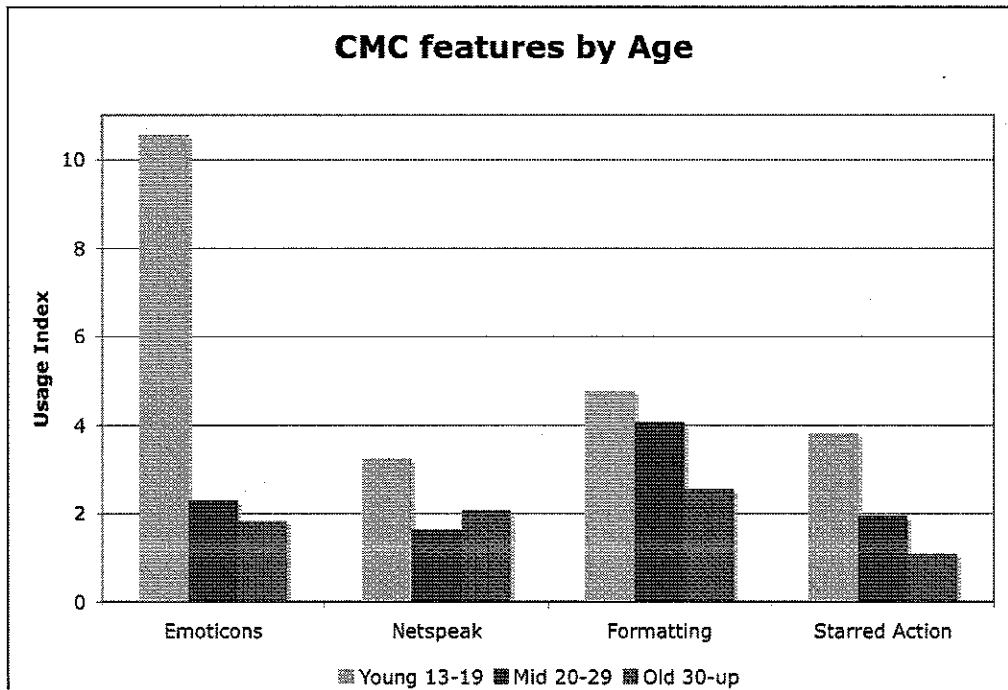
As can be seen from the figures above, both gender and age show clear trends of higher usage by females and by younger people. The one exception is a slight inversion between old and mid aged members for Netspeak, which is a less clear case because it violates the rules. This indicates that females and young members are better at adapting to CMC than males and older members.

In terms of gender, this finding indicates that findings about gendered language in real life have a counterpart online. Sociolinguistic investigations into real life speech communities have found that females tend to be more flexible in their style than males (Wardhaugh 2002:328). Wardhaugh relates this to the fact that "women appear to be more conscious of uses of language which they associate with their 'betters' in society, that is, those they regard as being socially superior," and relates this to the middle class (2002:328). However, in this instance we see females actually moving for a change away

from upper middle class standard English, and towards a new kind of computer mediated communicational standard.

Why is this the case? There have been many explanations put forth for why cross-linguistically females have more flexible language than males, ranging from psychological and sociological to biological explanations. What communication in the CMC environment can add to this debate is the fact that gender is not an immediately visible factor. Unless a member has a name that is clearly indicative of their gender (such as *paperflowergirl* or *Nick*), gender is not immediately apparent. Most of the posts were made during a period when the only way to find out someone's gender was to look it up on people's profile page. A couple of months before the sample ended, a small icon has been added to each member's avatars to indicate whether they are male or female if they choose to share such information. Still, females score higher than males.

This must then not be attributed to perceptions of gender in the conversation context, but rather to influences of such having produced higher awareness of language variation in females, which is then applied to CMC.

The other result derived from this comparison is that age is also an important factor. Older members consistently used less CMC features than younger members. This can partially be explained by the fact that the older members of the sample (those over 30) did very probably not grow up with CMC, but were introduced to it later in life. Online communities were formed as early as the late 1980s and early 1990s, but participation levels were much lower than they are now. In contrast, many more young people today take e-mail, instant messenger, and message boards as a given. Thus, they are more probable to be 'native' CMC speakers. While CMC features do not constitute a

language in the same way as English, exposure during younger years makes it easier to acquire an effective grasp of such features, and it is probable that some members came to the Leaky Lounge already familiar with them.

The fact that the same pattern of relative scores was produced for text formatting merits special mention. Unlike the other three CMC features investigated, text formatting is an established feature of traditional forms of written language. With the claims that CMC, or Netspeak, is sloppy language by some, and at least detached from standard constraints, by others, one would then have expected those who have adapted the least to Netspeak (i.e. older members) to use this feature the most. However, this is not the case. The best solution to this puzzle is to propose that CMC is not a language variation necessarily opposed to traditional written language, nor located somewhere between written and spoken language. It is a form in itself.

Bolds and italics, for example, fill the functions of traditional writing, as seen in the examples below, to mark titles, talk about words, and add emphasis.

22. If Tolkien were still alive, do you think the world would beat a path to his door and implore him to write a sequel to *Lord of the Rings*? You bet!
23. But word in the prophecy is *vanquish*.
24. And given that he **does** have that sense of honour, skewed though it may be, I truly believe that Snape would've told DD of Wormtail's deceit had he known about it.

But these formatting features, bold especially, are also used to draw attention to the mention of a specific name. Because there are multiple participants in most conversations (threads) this serves to clarify the addressee in the first example. Both examples show statements by members who do not bother to capitalize their *I*s, but who do take the time to type up the code to make their fellow members' name appear bold, showing not necessarily a need for quick expressions, as has been presumed to motivate

63

chat-room abbreviations, but rather a reprioritization of what information is important and what is not important to encode.

25. **HAngel**: I would love to join your and Thunderbird's conversation, but i am out of the loop right now. I need to catch up.
26. All in all, i think **Leo's** theory is the best one i've heard so far.

4.6 Harry Potter Speech Community Markers

In RL speech communities the linguistic markers particular to that community are often phonological in nature, such as in Milroy's study of Belfast dialects (Milroy 1980). However, for online speech communities such subtle differences in spoken language do not feature significantly in CMC, at least not currently, because spoken language is a small portion of the language being used. In the case of the HP online speech community, the fact that at least a significant portion of the members are non-native speakers of English would prevent the effective spread of such markers. However, there are other markers that do exist in CMC, even some that only exist in CMC. The markers examined here fall into the categories of lexical markers, abbreviations, constructions, expressions, and socially conscious usage. For some of these markers, empirical investigations were made to see whether their use was related to increased integration into the speech community.

In choosing what markers to pay attention to, and how to analyze them, it is important to distinguish between the particular language of the HP books, and the language of the HP speech community. This is in order to differentiate between those who have just read the books and therefore understand the meaning of terms such as *Patronus* and *Horcrux*, and those who are part of a group whose linguistic distinction is not merely derived from the *Harry Potter* books.

4.6.1 Lexical Markers

A certain scope of descriptive vocabulary has emerged among fans to describe the HP universe in ways other than those of J.K. Rowling, and to describe fandom itself. The most noteworthy terms are those that center around *shipping*. This phenomenon is not unique to Harry Potter fandom, but this is the place in which it is most infamous and elaborate in its form (Wikipedia "Shipping (fandom)"). By *ship* is meant a romantic relationship between two characters or RL people. By extension, those who engage in debates about 'who will end up with whom' are called *shippers*, or they *sail* a ship, e.g. Ron and Hermione. Ships can be *sunk* and *torpedoed* by good arguments, or by the definitive word of the author. The following sentences show how *ship* and the connected paradigms are used in context. The second example shows its usage at the most elaborate, in which a senior staff member of The Leaky Cauldron reprimands the participants in a discussion for their behavior. Because shipping is known to easily deteriorate into flames and flame wars, TLC has a policy of limiting the places in which such discussions should take place.

27. I think this also explains why some people ship Harry/Hermione so fanatically.
28. We do request that the Good Ship R/H and the Cruise Ship H/H please not torpedo each other in Leaky waters, and that the passengers of these two fine ships not spend too much of our bandwidth arguing over which boy Hermione thinks is the cutest - because until JKR finishes the series, it's an un-winable argument.[16]

The particular habit of referring to ships as naval ships is unique to HP fandom. In the above example, *Good Ship R/H* means 'Ron and Hermione' while *Cruise Ship H/H* means 'Harry and Hermione.' Alternatively, these ships may be referred to by another set

---

[16] http://www.the-leaky-cauldron.org/MTarchives/004452.html

of names, such as *HMS Pumpkin Pie* for 'Harry and Hermione,' or *S.S. Fire & Ice* for

'Draco and Ginny,' explained by a shipper below. These more 'mysterious' titles are

often based on the perceived qualities of the characters (Ginny being a warm/passionate

character and Draco being cold), but sometimes the name is completely irreducible.

Pumpkin Pie originated in a popular and rather explicit work of fan fiction. This kind of

titles are preceded by the initials used for real ships (for source, please see footnote[17])

29. The ship which I actively sail is the S.S Fire & Ice, which hopefully tells you who
we're shipping [Draco/Ginny]. The ship name came from the combination of both of
their physical looks, temperment, and Robert Frost's famous poem Fire and Ice. I'm
not sure who came up with the name, but we sure are indebted to him/her.[18]

The reference to *ships* is universal throughout HP fandom, even on sites whose

primary language is not English. However, it is probable that the term did in fact not

originate in HP fandom, but rather in X-Files fandom, where fans used the word *ship* as

an abbreviation when discussing the possibility of Mulder and Scully having a romantic

relationship.[19] Once taken up in HP fandom, the metaphor has been considerably

extended and popularized. As mentioned earlier, there is most likely a considerable

membership overlap between various online fandoms, which would account for the

integration of this term into the HP speech community.

---

[17] This information was taken from <http://www.lumosdissendium.org/dictionary.html> October
30th 2005. Examples were from the following thread:
<http://www.fictionalley.org/fictionalleypark/forums/showthread.php?s=&threadid=73709>
accessed October 30th 2005. For a list of all ships and their names, please see
http://forums.fictionalley.org/park/showthread.php?&threadid=4591, Accessed December 1st
2005.
[18] <http://www.fictionalley.org/fictionalleypark/forums/showthread.php?s=&threadid=73709>
accessed October 30th 2005
[19] From http://www.lumosdissendium.org/dictionary.html. This source is a fan site, and as such is
of limited reliability.

Shipping itself is very popular, but it is also frowned upon. A meta-discussion of the shipping phenomenon on the forum showed that the general presumption was that shippers were predominantly younger, and predominantly female.[20]

There are also other lexical markers, though their paradigms are not so complex as the *shipping* paradigm. One term that is unique to HP fandom is the concept of *fanon*, which refers to the collective body of fandom assumptions and theories about the HP universe and its characters, but which are not *canon*, i.e. the word of the author in books or interviews. Other word appropriated or invented for the descriptive purposes of HP fandom include the standard English word *spoiler* (information about yet to be released books or movies) and its non-standard derivative *spoilerific* (adj.). There are also several terms using *Potter-* as a prefix, such as *Potterhead* (hard core HP fan), *Potterverse* (fictional universe of Harry Potter), *Potterdom* (HP fandom), and *Potterwars* (copyright conflict with Warner Brothers over fans' right to have HP websites).

The results of the mining process for HP vocabulary showed, after some adjustments, a general increase in such terms. However, the initial data was overwhelmingly dominated by one member, whose usage index was about 20 times the average for the first 150 posts. Upon closer investigation, it became clear that this member was almost exclusively involved in shipping discussions during her earlier posts. Given that many of the HP fandom markers do have their origin in shipping, her intense involvement skewed the sample considerably, illustrating how context-dependent such lexical markers are. After this member was eliminated from the sample, the following distribution was produced:

---

[20] <http://www.leakylounge.com/cgi-bin/ikonboard.cgi?s=3746898c0f9917ec2cf29c9858b27efe;act=ST;f=65;t=2380;&#top> Accessed October 30th 2005.

Figure 20: HP Vocabulary
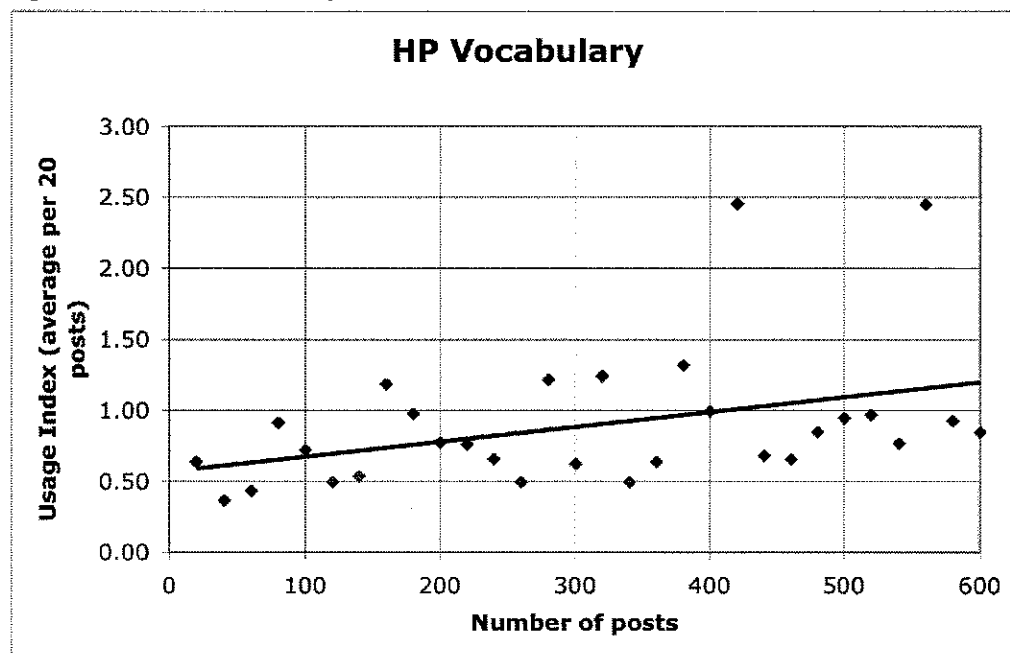
There is a gradual increase in usage over time. This may be indicative of increased involvement in shipping debates as members gain seniority. It may also, however, reflect an increased awareness of certain terms, such as *fanon* and *Potterverse*, whose usage is appropriate for most, if not all *HP*-related topics of conversation. Excepting the one abnormal speaker, the results showed no significant difference in usage between males and females, older and younger members.

4.6.2 HP Abbreviations

The extensive use of abbreviations is very common in online communication, particularly in synchronous chat, but also in asynchronous message boards (Crystal 2004;84). This is also the case in the HP online speech community, where one finds abbreviations of three sorts: those used quite widely in CMC, those pertaining to fan

fiction, and a whole range of abbreviations for elements from the HP books and for fandom concepts.

However, unlike some other speech communities, there is considerable tension in the use of abbreviations in several places in HP fandom, because they have a standard English policy intended to benefit new members and members who are not entirely fluent in English. Thus grammar mistakes are tolerated when honest, but "IMHO the HBP is DD"[21] would not be acceptable. The tension is between forces of habit and the convenience of shorthand on the one hand, and the desire to comply with the rules on the other.

Figure 21: HP Abbreviations
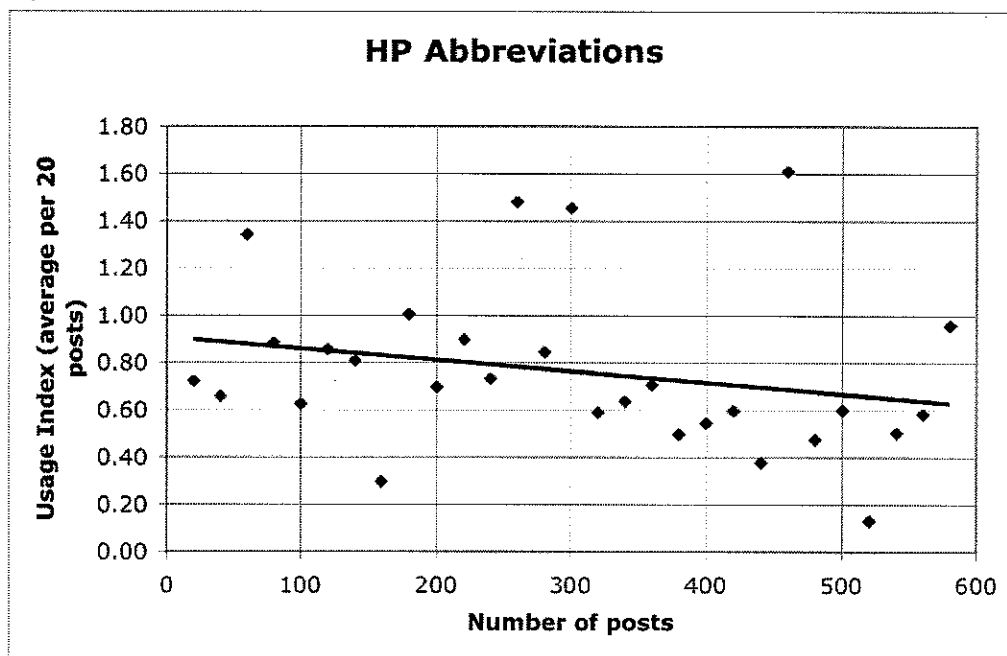
Changes in the use of general CMC abbreviations and shorthand expressions were investigated under CMC and Netspeak above. The results for HP abbreviations were somewhat puzzling, showing a decrease in usage over time. There appears to be no

---

[21] 'In my honest opinion, Dumbledore is the Half-Blood Prince.' *Harry Potter and the Half-Blood Prince* was the title of the sixth book, and there was much speculation about the identity of the title character.

reason for this, because HP abbreviations, while asked to stay within reason, are not explicitly limited by the forum rules, nor has the author perceived any negative stigma attached to the use of such abbreviations. The data is very scattered, and like HP-vocabulary, also very dependent on the topic of conversation. This is because most of the abbreviations are shorthand for fictional people and places, and the titles of the books.

30. As to whether or not TR was born evil, JKR has said that she doesn't believe anyone is born evil.
31. I wonder how it slipped past DD that Quirrel was hiding Voldie's head inside his turban?

As can be seen from the above examples, usage of abbreviations are very context dependent. *TR* will not appear in the conversation unless a member wishes to talk about Tom Riddle, and the same is the case for *Voldie* (Lord Voldemort), *DD* (Dumbledore), and *JKR* (Joanne Kathleen Rowling). The fact that frequency decreased over time may indicate that over time members spread their participation to include off topic threads not directly (or at all) related to the *Harry Potter* books and phenomenon, though more investigations would be necessary to validate such a claim. There were no significant age and gender differences in the use of HP abbreviations.

4.6.3 Constructions

There are a few productive constructions used in HP fandom. One is the use of a slash between the names (or initials) of two characters in a romantic relationship, as seen in sentence (1) and (2). This term is not confined to the HP speech community, but is an integral part of the shipping language.

32. I have always been a H/G kind of girl.

70

I have always been a 'Harry in a romantic relationship with Ginny' kind of girl.
33. X   /   Y
     Person X in a romantic relationship with   person Y

The second construction is of a less intuitive nature. It originated in fan fiction,

and is used to prevent confusions by modifying a character in the form

modifier!character, e.g. *bad!Harry*, which would mean a character who was just like

Harry, except that instead of being a generally good-natured boy, he now had a much

stronger bad side. Alternatively, this form could also refer to Harry's bad side. This term

probably has its origins in computer code, where the exclamation point was used to

denote *not*. The *not* is still implied because Pink!Hermione (Hermione wearing a pink

sweater in the Prisoner of Azkaban movie) is not Hermione from canon.[22] Two examples

of this construction used in context can be seen below. The first example illustrates its

inherent link to fan fiction. The person is commenting on the use of *S.S. Leather and*

Libraries as the title for the Draco/Hermione ship. The second example is not written in

the context of fan fiction, but is referring to a situation in which the character Barthemius

Crouch Jr. impersonated Alastor Moody through the use of a special potion.

34. S.S. Leather and Libraries
     This stems from Hermione's love of libraries and the Draco-Trilogy-inspired[23]
     LeatherPants!Draco.
35. I guess the whole Crouch!Moody thing has taught us not to take anybody at face
     value and that JKR is brilliant in writing characters that are "dual" in their motives.

4.6.4 Expressions

In the HP speech community there has been a considerable adoption of the

language and metaphors of the books into daily language. While some of these

---

[22] From http://www.lumosdissendium.org/dictionary.html. The source for this is a fan site, and as
such is of limited reliability.
[23] The *Draco-Trilogy* is presumably a popular fan fiction.

expressions derive their origins from the books, they are used in new contexts. For example the phrase "my two cents" is more often seen as "my two knuts" (a Knut being the smallest monetary currency of the Wizarding world). Examples of spontaneous metaphors using the language of the books in a new context are shown below,

36. There must be Death Eaters in the database eating your posts LOL

Death Eaters are the followers of Lord Voldemort in the HP series, and they are very bad people. This remark was in response to the complaint by the Leaky Lounge forum administrator that there is a problem with the database which is somehow preventing certain kinds of entries from saving, but that he cannot find or fix the problem. This comment jokes that maybe there are Death Eaters in the database who are causing the problem, adding 'laughs out loud,' in acknowledgement that this is humorous.

37. How rude! we should dungbomb them into oblivion!

In the HP books, Dungbombs are prank equipment that smell very bad when they go off. This is an expression of anger at abusive attacks by certain members on a forum moderator, but it is not intended to be acted on literally. Both these examples show how the make believe mindset of the forum creates new linguistic opportunities for expression and for humor. This will be discussed further in section 4.8 on the fantasy theme, below.

4.6.5 Socially conscious usage

Despite the appearance of being magical children's books, the *Harry Potter* series discusses some very serious themes though in a fictional setting. Consequently, there are bad characters in the books who use very derogatory language. The worst is the term

Mudblood, which is portrayed as just as offensive as contemporary RL racial slurs. What

is interesting to observe is the sensitive attitude of many members to this word.

38. Draco Malfoy hates "mudbloods", but probably not because of having been enraged
    by muggles, but because he's been brought up like this.
39. The Ministry knows that LV and the DEs target "Mudbloods",

The use of quotes in the above example illustrates the fact that some members are

uncomfortable using the term *mudblood* directly, and generally prefer the neutral terms

Muggle-born or *Half-blood* when referring to people's magical heritage.

Concluding the discussion of HP speech community linguistic markers, one can

thus see that there are many ways in which the HP speech community has adapted and

changed traditional language, the language of the *Harry Potter* books, and the language

of other online communities to suit its own expressive needs. Of the investigated

markers, the use of lexical (word) markers increased over time. However, much can also

be learned from a qualitative analysis of the ways in which members of the speech

community invent constructions and expressions that show their understanding of the

books and mark their membership the community of fans.

4.7 Norms of the HP Speech Community

Part of what defines a speech community is that they share communicative norms.

At the Leaky Lounge, there appear to be a few such norms, some of which are defined by

the rules, and others that are not. The most important rule at the Leaky Lounge is "Be

Nice!" While the moderators do work to ensure that this rule is upheld, the membership

itself is also very conscious of this rule, and the overall level of flaming and insults is

quite low, given the size of the Forum. The existence of this rule can be seen through the various ways in which the members seek to make their CMC friendly, such as through emoticons.

Many of the other norms relate to how one engages in the discussion. It is poor manners to post without first reading at least a few pages of posts, and when new members join a long conversation they tend to apologize for not having read everything, or 'brag' about having read the whole thread.

40. Ok, sorry to keep doing this to you... changing subjects, but this is a pretty open thread to throw in questions and ideas...
41. Forgive me for taking your topic on a slight detour

Also, members try to stay on topic, and tend to apologize when going on tangents or consciously try to change the topic of conversation. References to other works, perhaps especially Lord of the Rings and Star Wars, are often followed by an apology. This may be part of an attempt at differentiation from two of the fandoms most similar to that of *Harry Potter*, which is paradoxical because the scattered references to these works, and the responses, show that many, if not most, of the members of the Leaky Lounge are quite familiar with both Tolkien's *Lord of the Rings*, Peter Jacksons movie adaptation, and George Lucas' two *Star Wars* trilogies.

42. Perhaps he is like Anakin Skywalker / Darth Vader? (Sorry, just got back from Episode III and I cannot avoid making the connection [winking emoticon])
43. My pet theory (meow) is that Dumbledore gave a powerful Slytherin and a standout Gryffindor a love potion, in order to make a little Potter who would bring "balance to the Force" and finally get the factions to co-operate!
ducks tomatoes
Well, it could happen! :p

One norm to be mentioned briefly pertains to the use of stars to indicate action. Such action can only refer to the speaker. That is, one cannot describe the (imaginary) action of other members. As can be seen from the discussion of the Fantasy Theme shared by HP online speech community members (below), however, members maneuver around this prohibition by describing their own response to the anticipated actions of others. This effect can also be seen in the above example, where the member *ducks tomatoes* in anticipation of the displeasure of her fellow posters. It is unclear whether the rather unrealistic theory or the Star Wars reference are the cause of anticipated displeasure, but it is probably both.

44. I apologize for the immense length of this post.
45. Sorry if I was a tad long-winded, there...

In terms of length, Wick found that members of Pednet tried to keep their messages brief, and would apologize for 'imposing' on the others when writing a lot (Wick 1997: 66). At the Leaky Lounge, members appear to have a different norm regarding length: say something in each post, and take the necessary space. Thus, posts that are too short (one or two words) or posts that are very long, are stigmatized. Very short posts (called spam) are removed by staff, but members themselves apologize for very long posts (usually once they reach 1000 or more words).

4.8 The Fantasy Theme

At the Leaky Lounge, and other places in HP fandom, members appear to create a fantasy world. This phenomenon, described by Wick as a fantasy theme, a form of dramatizing communication, was also present on Pednet, where members engaged in a

pedestrian revenge fantasy (on drivers) (Wick 1997: 71). On Pednet this theme became a

lasting feature of the communication, invoked even after the original conversation died

(Wick 1997:71). On Leaky Lounge, this fantasy theme is so common and so diverse, that

it is perhaps more appropriate to call it a fantasy world, which in some ways 'merges'

with the CMC medium.

Why this is so common is hard to tell. The phenomenon might have its origin in

MUDs) which are online virtual worlds consisting only of text, in which members behave

as if they were in a fantasy landscape, much like conventional role playing games.

However, message boards are different because it is not a game, and members do not act

by proxy of a character. They do sometimes alter their self-representation to take on an

online 'identity,' such as that of a bird, but they do not pretend to be someone else, just a

variation of themselves.

Common to several of the modes of communication used by HP fandom members

is a tendency to extend the universe of the HP books (the *wizarding world*) to their

written and spoken CMC, and also to real life events. Thus, most web sites and forums

have a Harry Potter theme, such as Hogwarts, or Diagon Alley, conferences will sort

participants into Hogwarts Houses, and so forth. What this fantasy theme consists in is a

merge between the fictional reality of the books with virtual space. This is reinforced by

the fact that the Leaky Lounge structure matches the 'geography' of Diagon Alley, a

place in the Harry Potter books, and the various sub-sections of the forum are named after

known (or inferred) stores and places of the wizarding shopping district. News are

discussed in the Daily Prophet Offices, books are discussed in Flourish and Blotts, and

media discussions take place in the Wizarding Wireless Network sub-forum. The

fictional influence on this fantasy theme can be seen in the below examples, where

fictional characters are defined as part of the collective *we*, members wish characters

happy birthday, and allude to their own magical status.

46. He's already caused a death which we all took pretty hard ("we" being some people in the wizarding world and the readers of HP).
47. Also Happy birthday Ginny!!! [birthday emoticon] I hope you're celebrating with Harry :)
48. - Nelly (who often pretends she has not the Sight... :p)

Members also merge life on the screen with real life action, producing statements

that imply that the discussion thread and the forum in general are not virtually negotiated.

The first example presumes that the writing and reading of her message is in fact

synchronous (located in real time), though there might be a considerable lag before

anyone reads her post.   The double colons in the second example are a different notation

for the starred action discussed earlier.

49. Ok, I'm gonna quit here before we all fall asleep.
50. I'm celebrating that I'm the 50th poster in the LL!!!! I would like to thank ... All you guys for giving me someone to talk to... and JKR for writing these wonderful books... thank you guys!!! ::blows kisses while doing a Miss America type wave::

This fantasy theme, however, is taken much further than that found by Wick,

because the members also interact with each other and pretend to be in the same location.

They also gain a new level of detachment from reality. The revenge-scenarios Wick

found were sometimes far fetched, but never conflicted with the limits of the physical

universe, nor did they speakers situate their own location in a virtual space. The below

examples show members following each other around (in virtual space) by means of a

metal detector, and two others attending a party.

51. (psst: sorry I'm following you around...I've got a metal detector, and you've apparently got some galleons in your pocket...) :p
52. >500 posts! A round of butterbeer for everyone!
    What you say??? I can't here you over the music.

Sometimes all of these features are combined to create exchanges that loose most, if not all, of their reference to the real world. The following exchange is further complicated by the fact that Lord Muffinman is a moderator, and as such can write his own comments directly into the other member's post.

53. [Post by Nelly]
    Great Lord Muffything...
    salaams
    ...you are intelligenter than we...   :D [grinning emoticon]
    peeks
    but you don't happen to know anyone with the emotional range of a teaspoon, do you?   ;) [winking emoticon]
    Ouch! O my Master! Do not hex me! Ow! Please do not take it personally!
    Aieeee...

    (Lord Muffinman sez:) How else am I supposed to take it, worm!? Crucio![24]

    [Nelly again]:p [teasing emoticon] okay, don't kill me...

There are thus several features of CMC, some of which can be seen as direct extensions or proxy representations of real life communication. Others, especially when combined (as in the above example) create exchanges that are only possible in CMC. The closest RL equivalent would be actors on a stage with an infinite array of props, though even this would not convey the 'effective' use of magic by the participants.

5. Suggestions for Improvement and Future Research

The research conducted in this study can be further continued in two ways. The first pertains to the empirical aspects of the analysis. The sample employed here was

---

[24] Unforgivable torture curse from the *Harry Potter* novels.

quite small, and it would be reassuring, as well as interesting, to see how a larger sample, which ought to make the trends more explicit in cases of unclear results, supported the results.

One line of inquiry not pursued by this study, is to look at how different posting frequency and post length might impact the linguistic behavior of the poster. Certain correlations were found for age and gender, but are there other, virtual, demographics that have equal, or perhaps stronger, influence? That is, are prolific posters more community reinforcing than infrequent posters? Alternatively, is this more a function of length?

One thing that became clear as the study progressed was the close relationship between advanced use of computer mediated expressive opportunities, the fantasy theme, and tentatively, humor. If time had allowed, I would have looked deeper into the literature on humor and community building in real life, and sought to apply this to the use of humor in virtual community building. However, such an investigation will be for the future.

6. Conclusion

This paper has thus sought to situate the analysis of online speech communities in the wider context of speech communities and online communication. It has sought to supplement the rather scarce analysis of speech communities on the web, by learning from the work done by Cassell and Tversky (2005) and Wick (1997), and examining whether and how their methods and findings were applicable for the larger and more open speech communities to be found online.

Through the examination of Harry Potter fandom online, this study sought to explore the ways in which this speech community was being reinforced by the member's linguistic behavior. The markers of such behavior ranged from the examination of pronoun usage, various modes of interaction, CMC proficiency, community specific lexical markers, to norms and shared fantasy. The quantitative analysis established that most of the community reinforcing behavior increased as members gained in membership seniority. Some of these experienced a temporary decrease as novice members prove themselves as *HP-sleuths*. However, the overall trend was, as expected, an increase. Other analyses showed that age and gender sometimes played a significant role in impacting linguistic behavior online.

The qualitative analysis explored some of the ways in which members reinforced their community not easily captured by empirical analysis. The results showed that there are many ways in which members do this, both through conventional linguistic expressions such as greetings, but also several particular to the online medium, such as quotes, emoticons, and starred action expressions. The latter were found to also play an important role for the fantasy theme being acted out (virtually) by members of the speech community in order to increase their range of expression, but also for the purposes of humor, and tentatively also for community reinforcement.

7. Bibliography

ALEXA. Alexa Traffic Rankings. http://www.alexa.com Accessed December 5, 2005.

BBC NEWS, July 14 2005. The Magic of Selling Harry Potter. <http://news.bbc.co.uk/2/hi/business/4679553.stm>

BBC NEWS. October 4 2005. Global Potter sales top 300m mark.
    <http://news.bbc.co.uk/2/hi/entertainment/4308548.stm>

CASSELL, J. & TVERSKY, D. 2005. The Language of Online Intercultural Community
    Formation. *Journal of Computer Mediated Communication, 10* (2), Article 3.
    <http://jcmc.indiana.edu/vol10/issue2/cassell.html> Accessed September 26,
    2005)

CHAMBERS, J.K. 2003. Sociolinguistic Theory. MA, USA: Blackwell Publishing.

COLBERT, DAVID. 2004. *The Magical Worlds of Harry Potter.* New York, USA:
    Berkley Books.

. CRYSTAL, DAVID. 2001. *Language and the Internet.* Cambridge, UK: Cambridge
    University Press.

FICTIONALLEY. 2005. SCUSA Overview.
    <http://forums.fictionalley.org/park/showthread.php?&threadid=4591>
    Accessed December 1, 2005.

GRANGER, JOHN. 2004. *Looking for God in Harry Potter.* USA: SaltRiver, Tyndale
    House Publishers, Inc.

GREENFIELD, P.M. and SUBRAHMANYAM, K. 2003. Online discourse in a teen
    chatroom:New codes and new modes of coherence in a visual medium. *Journal
    of Applied Developmental Psychology,* Volume 24 (6).

JONES, QUENTIN. 1997. Virtual-Communities, Virtual Settlements & Cyber-
    Archaeology: A Theoretical Outline. *Journal of Computer Mediated
    Communication, 3* (3). <http://jcmc.indiana.edu/vol3/issue3/jones.html>
    Accessed October 20 2005.

KRONZEK, ALLAN ZOLA & KRONZEK, ELIZABETH. 2001. *The Sorcerer's Companion.*
    New York, USA: Broadway Books.

THE LEAKY CAULDRON. 2005. Rules. http://www.leakylounge.com/cgi-
    bin/ikonboard.cgi?act=Pages&page=Rules. Accessed October 26 2005.

MILROY, LESLIE. 1980. *Language and Social Networks.* Oxford, UK: Blackwell
    Publisher.

PENNEBAKER, J.W. & STONE, L.D. 2003. Words of wisdom: Language use over the
    lifespan. *Journal of Personality and Social Psychology, 85* (2), 291-301.

PORTER, CONSTANCE ELISE. 2004. A Typology of Virtual Communities: A Multi-
    Disciplinary foundation for Future Research. *Journal of Computer Mediated*

Communication, 10 (1), Article 3.
<http://jcmc.indiana.edu/vol10/issue1/porter.html> Accessed 26 Sept. 2005.

THURLOW, C., LENGEL, L., & TOMIC, A. 2004. *Computer Mediated Communication – Social Interaction and the Internet.* London: Sage Publications Inc.

WARDHAUGH, RONALD. 2002. *An Introduction to Sociolinguistics.* USA: Blackwell Publishers.

WATERS, GALADRIEL, *ed.* 2004. *The Plot Thickens… Harry Potter Investigated by Fans for Fans.* USA: Wizarding World Press.

WICK, NANCY. 1997. *Speech Community in the Virtual World: The Case of one Listserv.* Unpublished Ph.D. Dissertation, Department of Speech Communication, University of Washington, USA.

WIKIPEDIA Harry Potter Fandom/RPGs
<http://en.wikipedia.org/wiki/Harry_Potter_fandom/RPGs> Accessed October 26 2005.

WIKPEDIA, Shipping (fandom).
<http://en.wikipedia.org/wiki/Shipping_%28fandom%29> Accessed October 30th 2005.

8. Appendix

8.1 Visual Overview of the Forum
8.1.1 Example Post

Your messenger (1 new)

Welcome!
To TLC's First Forum!

The Rules : FAQ : Member List : Search : Staff

Part of the
Floo Network: News - Lexicon - Quotes - Galleries - Shop

Leaky Lounge > TLC Staff Chambers > Essay Project Drafting Room > Monday Meetup Thread

Forum News: NYC Premiere Wrap-up

1 Active User(s) in this Topic:

>paperflowergirl

Pages: (7) [1] 2 3 4 5 6 ...

Search Thread!

[Track this topic :: Email this topic :: Print this topic]

Topic: Monday Meetup Thread, Checking in!

< Next Oldest Topic | Next Newest Topic >

paperflowergirl
Edit this member

Schmancy Scribbler

Moderator

Posts: 858
Location: In the Essay Project Drafting Room
Joined: Jan. 2005
Post in this thread: #1

Posted: Oct. 24 2005, 12:19 am [Link]

Good Morning!

Clears piles of papers and guideline drafts off the biggest drafting table, and dishes up with hot chocolate and cupcakes.

Help yourselves!

This is something I have been thinking of setting up for a while. They have had one in the Mod Suite, and it is intended for people to check in (on Mondays, surprise surprise) just to show that we are all around, and for important messages to be communicated.

It is also good for some munching, especially of home baked goods.

I am currently being killed by my thesis, but it is moving ahead, and I'll try to pop by here as often as possible.

Oh, and could you all pick out your favorite titles in the Title thread?

_____

Proud to be a Hufflepuff!
Proud to be a Hufflepuff!

Search!

Search!

Search!

Search!

Search!

Search!

More!

Shopping at The Cauldron Shop supports this forum!

Back to top

IP: [165.108.211.224] Report this post to a moderator

8.1.2 Example Profile Page

The Rules : FAQ : Member List : Search : Staff

Leaky Lounge » Viewing member profile

Forum News: NYC Premiere Wrap-up

» paperflowergirl

(Send Private Message · Add to Contacts)

Personal Photo

Avatar and Member Pips

Post Information

Last Post: paperflowergirl last posted in Essay Project Drafting Room in Essay discussion forums on Leaky Lounge on Dec. 05 2005, 7:21 pm
Last Profile Update: Nov. 20 2005, 9:52 pm
Post Average: 2.76 Posts Per Day
Total Posts: 858 [Show all posts]

Contact Details

Email Address: Send paperflowergirl an email
AOL Instant Messenger Username: *No Information Entered*
Skype Username: *No Information Entered*
MSN Messenger Username: *No Information Entered*
Yahoo! Messenger Username: *No Information Entered*

Search!

Search!

Search!

8.2 Search Strings

General Information

Within the Apple utility Terminal, two especially written programs were used:
magic_grep, which eliminated extraneous material from the html files, making sure the
results only reflected language in the member's posts, and magic_window, which put the
results together in groups, set to 20 posts. The magic suite was written by Sean Christ,
professor of linguistics at Swarthmore College.

Certain characters, particularly the apostrophe, is represented by another string (') in the html. A limited number of regular expressions were used to create more flexible searches: the period(.) replaces anything, , star (*) finds zero or more occurrences of the previous character, and I functioning as 'or' within the search string. When preceded by a backslash (\), the real occurrence (as opposed to the regular expression) of punctuation characters is indicated. The example member given is paperflowergirl (the author's username), who was not part of the sample.

Length
By words
magic_grep " " paperflowergirl.html I magic_window 20 I head -50

Pronouns
First Person Singular
magic_grep "I I I\,I I\.I i I i\,I i\.I me I me\,I me\.IMy I my I mine I mine\,I mine\.ImyselfII\&\#39\;dIi\&\#39\;dII\&\#39\;llIi\&\#39\;llII\&\#39\;mIi\&\#39\;mIIm "
paperflowergirl.html I magic_window 20 I head -50

First Person Plural
magic_grep "We I we I we\,I we\.IOurI our I our\,I our\.I ours\,I ours\.IWe\&\#39\;llIwe\&\#39\;llIWe\&\#39\;dIwe\&\#39\;dI We\&\#39\;reIwe\&\#39\;reIL et\&\#39\;sI let\&\#39\;s" paperflowergirl.html I magic_window 20 I head –50

Interaction
Second Person
I forgot to record… but this is probably it
magic_grep "You I you I you\,I you\.IYourIyourIyou\&\#39\;reIyou\&\#39\;dIyou\&\#39\;llIy\&\#39\;allIyall"
paperflowergirl.html I magic_window 20 I head –50

Agreement
Final: magic_grep "I agreeI i agreeII concurI i concurIYes\,IYes\.IYes\!I yes\,I yes\.I yes\!IYeahIyeahIYESIGood pointIgood pointIGOOD pointIgood ideaIGood ideaIGOOD ideaII also thinkII also agreeII aslo thinkII aslo agreeII also concurIialso thinkIi also agreeIi aslo thinkIi aslo agreeIi also concurII too thinkI i too thinkII too agreeIi too agreeII definitely agreeII definately agreeIi definitely agreeIi definately agree"
paperflowergirl.html I magic_window 20 I head –50

Disagreement

Final: magic_grep "I disagree|i disagree|I do not agree|i do not agree|i don\&\#39\;t agree|I don\&\#39\;t agree|I dont agree|i dont agree|I am not sure|i am not sure|I\&\#39\;m not sure|i\&\#39\;m not sure|Im not sure|I do not think|i do not think|I don\&\#39\;t think|i don\&\#39\;t think|I dont think|i dont think|No\,|No\.|No\!|NO| no\,| no\.| no\!|No |no way|absolutely not|I dunno...|that\&\#39\;s not true|point is not valid" paperflowergirl.html | magic_window 20 | head –50

Greetings
magic_grep "Hi |hi |HI |Hi\,|hi\,|HI\,|Hi\.|hi\.|HI\.|Hi\!|hi\!|HI\!|Hello|hello|HELLO|Hullo|hullo|Welcome|welcome|WELCOME|nice to see you|Nice to see you|Hey| hey|greetings|Greetings|GREETINGS" paperflowergirl.html | magic_window 20 | head -50

Wh-questions
magic_grep "What.*\?|what.*\?|Who.*\?|who.*\?|Where.*\?|where.*\?|When.*\?|when.*\?|Why.*\?|why.*\?|Wherefore.*\?|wherefore.*\?|Who.*\?|who.*\?" paperflowergirl.html | magic_window 20 | head –50

Question marks
magic_grep "\?" paperflowergirl.html | magic_window 20 | head –50

Friendliness and softeners
Apologies
magic_grep "I apologize|i apologize|I apologise|i apologise|I am sorry|i am sorry|I\&\#39\;m sorry|i\&\#39\;m sorry|Im sorry|Sorry but|I did not mean|i did not mean|I didn\&\#39\;t mean|i didn\&\#39\;t mean|I didnt mean|i didnt mean|I feel bad|i feel bad|I did not intend|i did not intend|I didn\&\#39\;t intend|i didn\&\#39\;t intend|I didnt intend|i didnt intend|my apologies|My apologies|excuse me|excuse my|my excuse|Oo*ps|ooo*ps|OOPS|my bad" paperflowergirl.html | magic_window 20 | head -50

Personal qualifiers
magic_grep "I think|i think|I do think|i do think|I also think|i also think|I feel|i feel|I do feel|i do feel|I also feel|i also feel|I honestly think|i honestly think|I honestly feel|i honestly feel|I don\&\#39\;t think|i don\&\#39\;t think|I do not think|i do not think|I dont think|i dont think|In my opinion|in my opinion|IMO| imo|Imo|In my humble opinion|in my humble opinion|In my honest opinion|in my honest opinion|IMHO|imho|Imho|I always thought|i always thought|I always imagined|i always imagined|My two cents|my two cents|My 2 cents|my 2 cents|My two .nuts|my two .nuts|My 2 .nuts|my 2 .nuts|I assume|i assume" paperflowergirl.html | magic_window 20 | head –50

CMC Medium
Emoticons
magic_grep "\:\-\)|\:\-\(|\:o\)|B\)|B\-\)|x\)|X\)|X\-\)|@@|\^o\^|\^\-\^|\!\-\-emo\&" paperflowergirl.html | magic_window 20 | head –50

Netspeak

magic_grep
"loll|LOL|Lol|LOl|LMFAO|lmfao|Lmfao|LMAO|Lmao|lmao|ROFL|Rofl|rofl|gr8|hhok|HH
OK|ROTF|Rotf|rotf|RUOK|Ruok|ruok|OMG|omg|IMO|imo|Imo|IMHO|imho|Imho|IMNS
HO|Imnsho|imnsho|BTW|Btw|btw|OT| ot\,| ot\.| ot\:|FYI|Fyi|
fyi|asap|Asap|ASAP|b4|B4|BBL|bbl|tafn|ttfn|tty|lfa4n|TNX|Tnx|tnx|THX|Thx|thx|TX|Tx|t
x|2b" paperflowergirl.html | magic_window 20 | head –50
(aided by Crystal 2004;**some page**)

Starred Action

magic_grep "*.**" paperflowergirl.html | magic_window 20 | head -50

Formatting

magic_grep
"|</u>|</i>|</B>|</U>|</I>|</s>|</S>|</color>|</COLOR>|</size>|</SIZE>||
</FONT>" paperflowergirl.html | magic_window 20 | head -50

Fandom Lexical markers

HP vocabulary

magic_grep
"Ship|ship|Sail|sail|Sink|sink|Torpedo|torpedo|Sink|sink|Sunk|sunk|Canon|canon|Fanon|fa
non|Spoil|SPOIL|spoil|Potterhead|potterhead|PotterHead|Potterverse|potterverse|PotterVer
se|Potterdom|potterdom|fic |fanfic|Fanfic|Fan fic|fan fic|fanart|Fanart|fan art|Fan
art|Slash|slash|fandom|Fandom| .V. | ..V.. " paperflowergirl.html | magic_window 20 |
head -50

HP abbreviations

Magic_grep
"HP|DD|D'D|SS|GG|HH|RR|RL|SB|PP|JP|LP|RB|LM|Voldie|BC|QQ|MM|TR|LV|BB|AD|
AD|F\&G|F\+G|MEM|RW|UV|VW|WT|TMR|C\!M|Harrymort|harrymort|HarryMort|Gin
nymort|ginnymort|GinnyMort|Quirrelmort|quirrelmort|QuirrelMort|quirrlelmort|quirrellm
ort|NHN|BB|FF|HE|MoM|MOM|DoM|DOM|MMAO|GH|MM|DA|dada|Dada|DE|GP|HB|
PT|HG|HoH|MWPP|QWC|TWT|WWN|AK|Aked|DIGS|HERON|OT|OOC|IC|HMS|USS|
HMAS|HMCS|RMS|HMNZS|PS|CoS|COS|CS| cos|Cos\,|Cos\.|Cos
|PoA|POA|poa|PA|GoF|gof|Gof|OotP|OOTP|OTP|OtP|ootp|Ootp|HBP|Hbp|hbp|FB|FBAW
TFT|FBaWtFT|QA|QtA|QTA|QTTA|QttA|JKR|TLC|MN|LL|HPANA|NQM|OBHWF|SC
USA|RPG" paperflowergirl.html | magic_window 20 | head -50

8.3 Overview over sites mentioned

CoSForums – www.cosforums.com
DarkMark Forums – www.darkmark.com/forums
Fanfiction.net – www.fanfiction.net
FictionAlley – www.fictionalley.org
Harry Latino – www.harrylatino.com
Harry Latino Foros – www.harrylatino.org
Harry Potter Automatic News Aggregator (HPANA) – www.hpana.com

Harry Potter Automatic News Aggregator (HPANA) Forums – www.hpana.com/forums
Harry Potter Lexicon – www.hp-lexicon.org
The Leaky Cauldron – www.the-leaky-cauldron.org
The Leaky Lounge – www.leakylounge.com
Mugglenet – www.mugglenet.com
Veritaserum – www.veritaserum.com

8.4 Acknowledgements