

A Brief Overview of the Linguistic Attributes of the Blogosphere

Stephanie Nilsson
D-Essay, Final Draft
Umeå Universitet
Engelska lingvistik

INTRODUCTION:	3
PART I--A SHORT HISTORY OF THE WEBLOG	4
THE WUNDERKAMMER: BLOGGING DEFINED	5
UNLEASHING THE POWER: BLOGGING TOOLS	8
PART II—ETHICAL ISSUES	9
WHO’S RIGHT IS IT, ANYWAY?	9
PART III—THE WEBLOG NETWORK	12
VIRTUAL COMMUNITY OR RANDOM INDIVIDUALS SHARING A COMMON INTEREST?	12
TO KNOW SOMEONE:	13
THE COMMUNITY—A VISUALIZATION	14
FIGURE 2:	14
FIGURE 3:	15
THE RESEARCH NETWORK—A LINGUISTIC ANALYSIS	19
THE HYPERLINK	20
IN-GROUP/OUT-GROUP LANGUAGE	23
PART IV—THE POST	24
THE BASIC POST—AN ANALYSIS OF BASIC LINGUISTIC ATTRIBUTES FOUND IN THE MAIN POSTING AREA.	24
PART V—BLOGSPEAK	27
NETSPEAK	27
WHAT THEN IS NETSPEAK?	29
BLOGS	29
WRITTEN LANGUAGE CRITERIA	29
THE WEB	29
BLOGS	29
BLOGSPEAK—THE LANGUAGE OF THE WEBLOG.	30

FROM THE SPEECH PARENT:	31
FROM THE WRITING PARENT:	32
ORAL AND LITERATE STRATEGIES:	34

CONCLUSION	37
-------------------	-----------

BIBLIOGRAPHY	38
---------------------	-----------

Introduction:

At first glance, a weblog appears to be the musings of a lone webcrawler, trolling the sea of information that is the World Wide Web. A closer look, however, will reveal a micro world of conversation and monologue joined in a seemingly endless net of hyperlinks. Weblogs are interesting in that they contain attributes of both the monologue and the dialogue. They are simultaneously self-reflecting journal entries and open-ended invitations for conversation. This unique dichotomy lends itself to the evolution of a new variety of language that has evolved from spoken and written communication, and adapted itself to flourish in the virtual environment.

The purpose of this study is two-fold. Firstly, this study will attempt to identify the structure of the weblogging community. A graphic representation has been created from empirical data collected from certain forms of hyperlinks. The community will be identified according to its structure and network strength. As certain types of networks suggest particular language norms, the structure determined here will be assessed according to the absence or presence of these norms. The second goal of this study is to determine if a separate variety of language exists in the weblogging community. This variety would be an extension of the popular term, *netspeak*¹, based on the distinctive form of discourse found within weblogs. For example, does conversation exist, and if so, how does it flourish in an environment where authors own and control their own space? Does the language of a weblog support group identity and strength? As this study was undertaken using the social network analysis approach, certain assumptions have been made. Namely, the nature of the virtual relationship is important as is the density of the network in that it directly relates to the potential for variety in communication.

This paper will be broken into five sections. The first section will define the weblog as a genre, giving a short history and explanation of its value as an on-line tool. The

¹ Netspeak, according to David Crystal is “a type of language displaying features that are unique to the Internet...arising out of its character as a medium which is electronic, global, and interactive” (Crystal, 18).

following section includes a discussion of the ethical guidelines employed to research the on-line community.

The third section will begin by defining the virtual community. It will go on to examine the notion of social networks, arguing that the network appears to be closed and dense, but retains an opening, or pathway that allows considerable movement in and out of the core group. The methods and procedures used to analyze the structure and function of the group, as well as graphical representations will be presented. The fourth section will analyze the language of the post, or the main discussion area of a weblog.

The final section will suggest the possibility of a new variety of language found within the weblogging communities. The archiving function of weblogs allows the researcher access to the natural language of the weblog author, and opens the door for the researcher to explore the possibility of the evolution of a new variety of language in the weblogging environment. The variety of language proposed in this paper originates from a combination of sources; the need to communicate, the nature of the on-line experience, as well as the linguistic variety, Netspeak. The weblog variety will be introduced within a metaphor describing itself as the daughter strain of the parent forms of communication; speech and writing. The reader should keep in mind that this section is based on a very limited study and any suggestions made will require further research.

The study of the structure and variety of weblogs is important as different kinds of social networks suggest different types of language maintenance. Groups that are close-knit exhibit capabilities to resisting outside forces, not only linguistically, but socially as well. These groups often create their own lexicon of jargon and taboo words that both mark member as either in-group or out-group, but also help strengthen the sense of group identity. Groups that are fairly open exhibit great fluctuation in variety between different linguistic situations.

Part I --A Short History of the Weblog

To me, the blog concept is about three things: Frequency, Brevity, and Personality. (..) This clarification has evolved over time, but I realized early on that what was significant about blogs was the format -- not the content. (Evan Williams, creator of Pyra)

The Weblog is as old as the Web itself (see Blood 2001). It began when computer savvy web designers and intelligent web enthusiasts wanted to record their wanderings though

Cyberspace. These individuals collected links and posted them on frequently updated sites that resemble what most know as a home web page, or its cousin the What's New page. Linked were pages that the web savvy thought others may be interested in, but had not found in their own on-line travels. Originally, pages were kept by those that knew how to design websites. Weblogs were somewhat difficult to maintain as one needed to code or use HTML each time the site was updated. Consequently, these pages often contained links with little to no commentary. These pages were used to filter the web, and were especially useful as the now-popular on-line search engines had not yet been developed. Aptly named, the filter weblog began to carve out the format which has now developed into a journal-style weblog that goes beyond the commented link to include commentary on a person's existence.

The Wunderkammer²: Blogging Defined

Many of the pioneers of weblogging have developed their own definitions of what constitutes a weblog, and often include its importance to the web. The weblog is a frequently updated website that lists entries in reverse chronological order. Often weblogs include a list of links to other weblogs and to sites that interest the author. They are usually highly specialized to a certain topic, or group of topics, written through an individual point of view. There are some weblogs written by groups of authors, but these authors usually write in a similar style and with similar points of view. The evolution of weblogs and the community that has developed around them has led to new features such as commenting functions, trackbacks, and interlinking. These features will be delved into deeper as community and networking is examined later in the paper. Rebecca Blood (2002), agrees with Evan Williams that the format is the defining characteristic of a weblog³. Blood goes on to state that a great benefit of keeping an on-line journal is the creation of a more critical thinker and a better writer. She believes that the act of filtering itself creates an opportunity for a writer and websurfer to hone his or her filtering skills or face the embarrassment of posting entries that appear shallow and uninteresting. "(..)every weblogger who maintains a filter quickly becomes aware that she is building her reputation with the links she provides, and as a result few people are willing to put just anything on their page (Blood 2002:31)." Jesse James Garrett, content editor for Ingram Micro's website defines the uniqueness of the weblog as such, "weblogs are the pirate radio stations of the Web, personal platforms through which

² A Wunderkammer was, literally, a cabinet of wonders. The forerunner of museums, it emphasized typology rather than chronology. For a more complete definition, see <http://linux.nscad.ns.ca/~bobr/wcab/whatis.html>.

³ See **Figure 1** for an example of the format of a weblog.

individuals broadcast their perspectives on...basically anything...that strikes their fancy from the vast sea of raw material available out there" (Katz 2002:22). Garrett defines a quality weblog as, "neither a faceless news-gathering organization nor an impersonal clipping service, a quality weblog is distinguished by the voice of its editor, and that editor's connection with his or her audience" (Katz 2002:22). Joe Clark, author of Joe Clark.Org, has described the weblog as "a form of exteriorized psychology. It's a part of you, or of your psyche; while a titanium hip joint or a pacemaker might bring technology *inside* the corporeal you, a weblog used technology to bring the psychological you *outside*..."(Clark 2002:68).

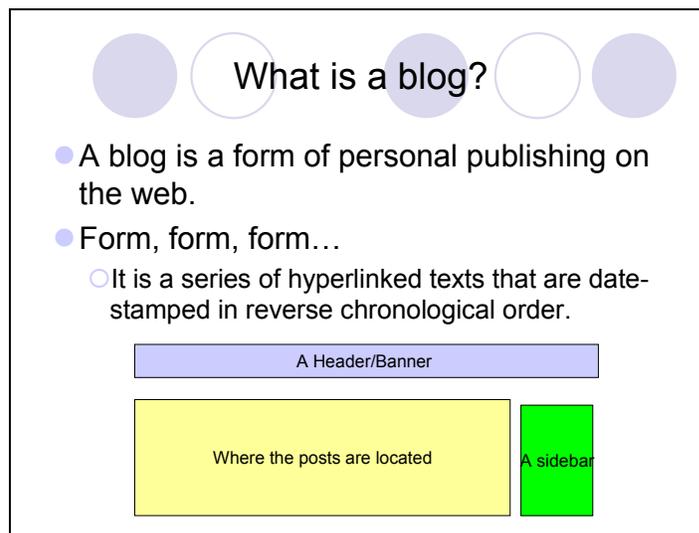


Figure 1- Slide from my HUMLab seminar (May, 2003), "att blogga"- Format

As the community began to develop, writers began to dislike the term "weblog" as it was seen as a boring name for an exciting new development in on-line publishing. Peter Merholz, author of the weblog PeterMe (<http://www.peterme.com>) stated in 1999 in his popular log that he was going to pronounce it "wee-blog". This name stuck and was eventually shortened to "blog" and the writer became, a "blogger" (Blood 2002:7).

A good label for a weblog, one that summarizes what many writers and editors have struggled to define, comes from Julian Dibbell in his description of Jorn Barger in *Portrait of the Blogger as a Young Man*. Dibbell likens a blog to a wunderkammer, a treasure trove of curiosities, capable of eliciting the mystique and wonder. Due to the filtering effect of old-style blogs, wonders can indeed be found within their many pages. The newer diary-style blog is less like a wunderkammer and more like a prism, through which the reader learns

to *know*⁴ the author through the multifaceted tones that lie just below the surface of the daily post. The regular reading/writing exchange between the reader and author makes apparent changes in tone and signal more than what the post may or may not say. An example of this occurred in the postings of one of the hubs⁵ of the research group examined in this paper. This person's posts suddenly became short and curt, not at all like the usual perky and lengthy posts previously uploaded to the blog. After a few days of this type of posting, the community surrounding this blogger began commenting, asking if everything was okay. Surprised at the community's astuteness, this blogger disclosed that they were going through a difficult breakup.

Many critics of weblogging have stated that those who keep weblogs are participating in a large game of mental masturbation. Critics accuse these writers, or bloggers as they are known as in the weblog circles, of using their weblogs as a place of exhibitionism, saying nothing significant and wasting time and space on the World Wide Web. Ben Brown, shouted affirmation for this view of weblogging in an open letter by in *Teeth* e-zine's that stated that "You're (a blogger) not a designer, you're not a writer, and you're not an editor!" (Dibbell 2002:72). Dibbell eloquently counters this claim stating:

"Well, no blogger, you're not. And therein lies your gift. Because even if it is true, the vast majority of blogs would not be missed by a handful of people were the Earth to open up and swallow them, and even if the best are still no substitute for the sustained attention of literary or journalistic works, it's also true that sustained attention is not what weblogs are about anyway. At their most interesting they embody something that exceeds attention, and transforms it: they are constructed from and pay implicit tribute to a peculiarly contemporary sort of wonder... a random collection of strange, compelling objects "(Dibbell 2002: 73).

Dibbell goes on to state that just as the *Wunderkammer* has evolved into what we now know as the museum, so will the weblog evolve into something that defines the way we exist in cyberspace. He also goes on to warn the keepers of weblogs to enjoy the art of blogging while it lasts because, as the wonder of a *Wunderkammer* was replaced by the cold mastery of science, so can the wonder of the World Wide Web be replaced by the cold mastery of cyber theory.

⁴ The concept of *knowing* is examined later in the paper.

⁵ Due to the personal nature of the posts, the name will not be disclosed.

Blogs are not about exhibiting your life in a public forum. Nor are they similar to static webpages, with only one opportunity to present an object/person/company in a favorable light. Those who regularly write a personal weblog know that blogs are subtle. It can be argued that the meat of a blog is not in the text or the links. The heart of a blog is located between the lines of random thought. It is hidden ever so carefully in the daily musings and consistency found in regular bloggers. Those that blog on a regular basis form a community, who on the surface appear to be a bunch of random links that occasionally mention each other. In reality, this community is comprised of faithful readers, concerned about the on-goings of each other's lives. They know by the tone of the post whether or not the prism is reflecting joy or sadness, hope or despair...and they in turn respond.

Unleashing the Power: Blogging Tools

The development of software that allows the blogger to blog without the tedious and time consuming need for HTML has led to rapid evolution in weblogs. Software like Pyra's Blogger has created an opportunity for ordinary people to blog without cost and with little effort. The new ease of blogging marks a freedom in publishing never before seen. In theory, it equalizes the playing field and allows anyone with a computer, Internet connection, and a yearning to be heard, the right to publish freely. Blood sees this shift as "a staggering shift from an age of carefully controlled information provided by sanctioned authorities (and artists) to an unprecedented opportunity for individual expression on a worldwide scale (Blood 2002: 15)." The blogger suddenly has the opportunity to say exactly what he or she thinks, without the worry of alienating readers and advertisers. In practice, however, the blogging sensation has reached such proportions that instead of weblogs filtering the web, a new format is needed to filter the immense amount of weblogs. This need is answered by the existence of links within sidebars. Nearly all weblogs are formatted with a main area of content and a sidebar with links. Regardless of the number of sidebars, the function stays somewhat static. These links play a double role in blogging. Importantly, these sidebar links provide a filter-function. They represent the bloggers' idea about which pages are worth reading and often, which pages are similar to his or her own. Secondly, these links are a marker of identification into which social network a blogger belongs. The notion of social networking will be examined further in the part three.

Before an in-depth discussion of social networking in cyberspace can be undertaken, it is first necessary to discuss the ethical implications behind such research. It is fundamental for researchers in the physical world to keep in mind the ethical needs of his or

her subject. It is just as imperative for researchers in the virtual world, as the security of anonymity is largely a myth. Part II of this paper will cover these issues in more detail.

Part II—Ethical Issues

Who's Right is it, Anyway?

The power to express oneself freely carries with it many questions common to Internet publishing. Namely, is the information found on a weblog public or private; what responsibilities do researchers have to the individual authors or blogs; who owns the information found in weblogs? These questions are currently hot topics of debate among both Internet researchers, as well as among the bloggers themselves.

A central problem that a researcher faces when conducting research on the Internet is the notion of public versus private domain. The fact that researchers do not need to identify themselves, nor identify their purpose within the anonymity of the Internet demands an answer to this question to protect the citizens of cyberspace. This problem was addressed at a 1999 workshop of the American Association for the Advancement of Science(AAAS) on ethical and legal aspects of human subject research on the Internet. In response, the AAAS developed three guidelines for the researcher concerned with on-line environments, autonomy, beneficence, and justice.

An individual should be treated with respect, and any available precaution should be taken to conceal his or her identity. Informed consent is usually required in real-world environments. In cyberspace, however, it is a little more difficult. Informed consent implies that the individual is aware of the consequences and benefits involved in the research and is a willing participant. Because of the fluid nature of the Internet, informed consent can be difficult, at best, to acquire. The division, between public and private, is central to determine when informed consent is required. Researchers that conduct their studies within a public forum are exempt from obtaining informed consent. The AAAS defines a public internet forum as a somewhat complicated thing. In the real world, data can be collected from public forums such as newspapers, television, or in public spaces. The nature of the Internet is easy public access to nearly everything, including seemingly private discussions, such as those found in forums of survivors of abuse, or those suffering from emotional problems. It is here that opinions split on the public vs. private debate. Because information is so readily accessible, does that make it public? The two schools of thought appear to be rather evenly

split. One group feels that if the website is publicly accessible, it is public. These researchers take the stand that it is up to the individual or moderator of the group to filter out what may be too personal to publish on the web. The other side of this argument is based on the fact that the Internet builds a false sense of anonymity, that people who post to these groups, such as support groups on-line, expect the same level of privacy that they would find in, for example, a real-world Alcoholics Anonymous meeting. Members of this group suggest that the “evaluation criteria for the level of sensitivity that members of a particular online community may expect, be proportional to the community’s level of *accessibility*” (AAAS 1999: 8)

Blogs fall into the public domain for several reasons, most importantly; blogs are a form of publishing. According to the Oxford On-Line Dictionary, to publish is to “To make public. I. 1. a. trans. To make publicly or generally known; to declare or report openly or publicly; to announce; to tell or noise abroad; also, to propagate, disseminate (a creed or system).” (Oxford 2003). Further proof of a blog as a form of publishing can be found in the name or description of blogging software. For example, www.blogger.com, the brainchild of Pyra, describes its function on the masthead of its website as push button publishing for the people. Movable type, currently the most popular weblogging tool, defines itself on its front page as “a decentralized, web-based personal publishing system designed to ease maintenance of regularly updated news or journal sites, like weblogs” (www.movabletype.org) Torill Mortensen, author of “thinking with my fingers”⁶ wrote this in an email about the public ness of blogging.

“Blogs belong to the public domain. They have been made public, and you as the writer and editor have full responsibility for what you write there. You can be quoted, but you should be properly cited when your thoughts and ideas are taken from your blog. Which means that having published something in your blog, you have actually published it, and can claim ownership to your ideas” (Mortensen, 2003-04-03 10:07).

Another feature of a blog that makes it distinctly public is its ability to archive. Archiving allows the public access to records years after the blogger has published his or her thoughts. It is obvious that blogs are highly accessible in the Internet. Blogs are not password protected, they are forms of on-line publishing, and they archive information for a theoretically infinite amount of time. How, then, does this level of accessibility measure proportionally to the members level of personal privacy.

⁶ <http://torillsin.blogspot.com/>

A large debate is raging though the blogging community about the amount of identity and personal information that should be included in blogs. Many bloggers feel that, although blogs are a public medium, they also hold a deeply personal link to the author. The question arises if blogs are publicly accessible works, or "electronically captured records of private conversation occurring in a quasi-spatial domain" (AAAS 1999:11).

The three criteria I used when establishing ethical guidelines for this paper were as follows:

1. How accessible are blogs?
2. To what degree is privacy perceived by the authors of blogs?
3. To what extent does personal identity exist in a blog?

To answer the first question, I looked at the ease of accessing blogs and their archives. As blogs are not password protected, and their archives are linked to the main page in an easily recognizable and accessible format, I determined that blogs are available to all.

I was able to gain a unique perspective on question number two through a blog "discussion" about trolls. A troll is a person that lurks in the background, reading posts, springing up occasionally to comment in rude and unconstructive ways. Jill Walker of [Jil/txtl](http://huminf.uib.no/~jill/archives/blog_theorising/a_troll_in_the_comments.html)⁷ announces the existence of a troll in her blog as, "I think I just got my first ever mean, bitchy and [unconstructive](#) comments. I'd actually thought that weblogs were fairly immune to that, and that it was mailing lists suffered from trolls". The discussion of trolls brought out the underlying discussion of privacy and rights to free speech over the Internet. The discussion centered around the question of whether or not to use the function of IP banning⁸ to block this particular person from posting anything more on Jill's website. Many bloggers were in favor of banning the troll, but Jill and a few other high profile bloggers eventually found the troll's comments to be thought provoking as one needed to take a stand and defend their point of view against the often inapt attacks from the troll. Also, the nature of the comment function in a blog allows for the reader to become a part of the blogger's script. The main portion, possibly 95% or more of the blog, is written by a single author. Commenting, on the other hand, takes a bit of the privacy away from the author and shares it with the audience, as was shown in the case of Jill's troll.

⁷ http://huminf.uib.no/~jill/archives/blog_theorising/a_troll_in_the_comments.html.

⁸ IP banning involves the computer generated ban of an IP address. If a person tries to log on to a certain website from a computer assigned with the banned IP address, he or she will be unable to access the site.

Personal identity, the main focus of the third and final ethical guideline in my research, is also a notion recently discussed by the blogging community. Linguistic research conducted on-line deals mainly with written, typed language, rather than spoken. This language leaves a physical record that may be examined by anyone. Because of this fact, I had to decide whether or not to attempt to conceal identities or to cite them. The AAAS reported that a participant in their conference stated that "from her experience in working with on-line communities that people invest just as much into their online identities as they do in their real ones" (AAAS 1999:13). The AAAS went on to state that, "it may not be enough to protect the real identities. It may be necessary to protect the online identities of the research subject as well" (AAAS 1999:13). In trying to reconcile the *publicness* of this publishing community with the personal identity embedded in their individual blogs, I have decided to assign pseudonyms to weblogs cited if the citation is used in such a way that harm may be inflicted upon the citeé.

PART III—The Weblog Network

Virtual community or random individuals sharing a common interest?

Do blogs make up a defined community, or are they merely detached mumblings from veteran webservers? This section of the paper will examine the notion of community in relation to blogs. Community, as defined by Milroy (1987:14), is a group that is cohesive and to which people consciously feel that they belong. Linguistically, this sense of cohesiveness can be examined from many different angles, depending on the dynamics of the group being observed. Blogs, and their unique format, provide the researcher with a special dynamic, the link. Links are multi-dimensional. They allow the surfer to glimpse a part of the poster's personality by linking to the poster's interests and hobbies. Links also allow bloggers to interact with each other. Brad Graham, author of Bradlands.com gives an example, "My weblog is linked from several others, and theirs from mine. We are a community, of sorts, a small town sharing gossip and news, recreation and sport, laughter and tears, all for the commonwealth" (Graham 2002:39). Ethan Eismann, a speaker at the first European BlogTalk

conference in 2003 in Vienna defined the blogging community as, “groups of authors committed to a knowledge domain who interact through social software to produce information artifacts that support shared goals and projects and sustain dynamic social networks.”

The sense of community will be measured in two ways. First, links will be measured to determine the connectedness of the community. The type of community determined from this data often suggests certain linguistic norms. The absence or presence of these norms will be explored through linguistic markers. This study will examine markers within two categories: structure and social behaviour. Within the category of structure, the rate of in-group versus out-group language will be measured. The category of social behaviour will measure instances of solidarity within a group as displayed by hyperlinks within the main posting area and the comments area.

To know someone:

Before a discussion about a virtual community can take place, the concept of knowing someone must be addressed. To know someone in the virtual world is different than knowing someone in the physical one. This concept, however, is very important to the theory of social networking. Knowing a person, or a group of people, in the real world usually involves contact with that person. You “know” someone when you have chatted with them several times, have been in contact with each other, or are otherwise acquainted with a person. You share a mutual trust and forms of communication. Knowledge of a person in the real-world often involves physical presence as well. Cyberspace is a little different, however. The element of *physical* presence is often missing in an on-line relationship, although *virtual* presence does exist. A blogger typically invests a lot of time writing his or her blog. As the writing progresses, the blogger’s personality is unveiled through his or her use of language. David Weinberger of johno.com at the Vienna Blogging Conference presented the theory that blogs allow the blogger to create a virtual self that is continuously present within the weblog. The notion that people have an inner/private self and an outer/public self is inconsequential in blogging. Weinberger contests that there is no inner self in a blog, that all blogs are written with an authentic public self. These on-line personas gather in what can be called the third

space⁹. This theory seems to be supported by the AAAS due to their suggestion to online researchers to protect the online identity of the individual as well as their physical one.

The Community—a visualization

Network strength can be difficult to establish during the linguistic evaluation of a community. Although it can prove to be invaluable as a diagnostic tool, the validity of network strength can be questionable. In order to acquire a better understanding of the type of community represented by university researchers using weblogs, graphic representations of their convergence have been created. The two diagrams below are of approximately the same group, with the exception of necessary linking differences, which will be explained in more detail for each individual picture below.

Figure 2:

Figure 2 is a graphical representation of seventy-five authors of topically related weblogs. Among these authors, 5625 links were examined for interrelatedness. These authors are all university researchers, primarily interested in the study of on-line environments. From each weblog, the sidebar links were taken and correlated into one spreadsheet with the names occurring down a list, as well as across the top of the spreadsheet. The seventy-five names were then cross checked for cases of interlinking or reciprocal linking. If a researcher was referenced (linked to), then that researcher was given a score of 1 under the name of the person who created the link. This crosschecking insured that all links were accounted for by the individual who created the links, as well as for the researcher who received a link. The spreadsheet was then run through Mage, a map creator found within the program, Ucinet 6 for Windows, (Borgatti, Everett, and Freeman).

Each line in figure 2 represents a hyperlink, the starting point being the weblog that created the link, and the end point the blog that received the link. As this diagram does not contain arrows, these points are impossible to detect, as well as unnecessary. The purpose of this diagram is not to determine who sent and who received a link, but rather the overall shape of the community. Each dot in the diagram represents a weblog. It is obvious from the diagram that there is a tight clustering of weblogs in the centre of the community. These weblogs give and receive the most hyperlinks and represent a closed dense network of

⁹ “a place where a number of people regularly go on a daily basis. It should be a place where the person feels welcome and comfortable, and where it is easy to enter into conversation. And a person who goes there should be able to expect to find both old and new friends each time she or he goes there” (Oldenberg, 1989).

blogging researchers. The weblogs located on the periphery of the cluster have a lower rate of linkage, yet continue to remain closely tied to the central cluster. An interesting feature of diagram 2 is the portal in the lower right hand corner. This outlet represents the links received by weblogs that are not in the condensed centre group. This outlet is unique in that it addresses a major problem found in dense communities, that of the stranger's role. The role of the stranger is unique in that acceptance into the community is quite hard to achieve, whether you have been a member of that community 5 minutes or 5 years. This outlet lets the stranger use the *friend of a friend* introduction into the network. This type of introduction allows the stranger instant access into the community through the trust previously built by the referrer. Through this portal, relatively new or unknown bloggers can receive a link from an *elite* or centrally located blogger and possibly increase his or her position in the community.

Figure 3:

Figure 3 was created with an on-line program called [blogstreet.com](http://www.blogstreet.com)¹⁰. In this program, the URL for the site in which one is interested is typed in to the space provided. The program automatically creates a visual neighbourhood based on the incoming and outgoing links of the URL in question. The author of the map may click on any hyperlinked hub to expand or contract the map. This map is not representative of the entire network of 75 researchers; it does, however clearly represent the principle of anchorage, a theory that a network can be considered from the point of view of the individual.

Each spike represents a linked blog, the blog being identified by the bulb at the end of the spike. At the centre of each cluster of spikes, the weblog responsible for the links is highlighted. This can be difficult to distinguish in the dense centre of the diagram, but is clearly illustrated in the outer blogs. Each weblog is at the centre of its own network, yet is simultaneously linked with the dense centre cluster of blogs that represent the most popular, or elite blogs in the social network. The popularity of the centre blogs is apparent due to the amount of links they receive. More about the significance of anchorage is discussed in the section entitled *The Research Network*.

It is important to note that both diagrams show a clear dense centre core with an equally clear periphery membership. These diagrams were produced using two different programs and with slightly different linking data. The similarity in the network pattern of this

¹⁰ <http://www.blogstreet.com/>

blogging group remains, however, static. This suggests that the emerging picture of a socially dense and closed linguistic network is accurate.

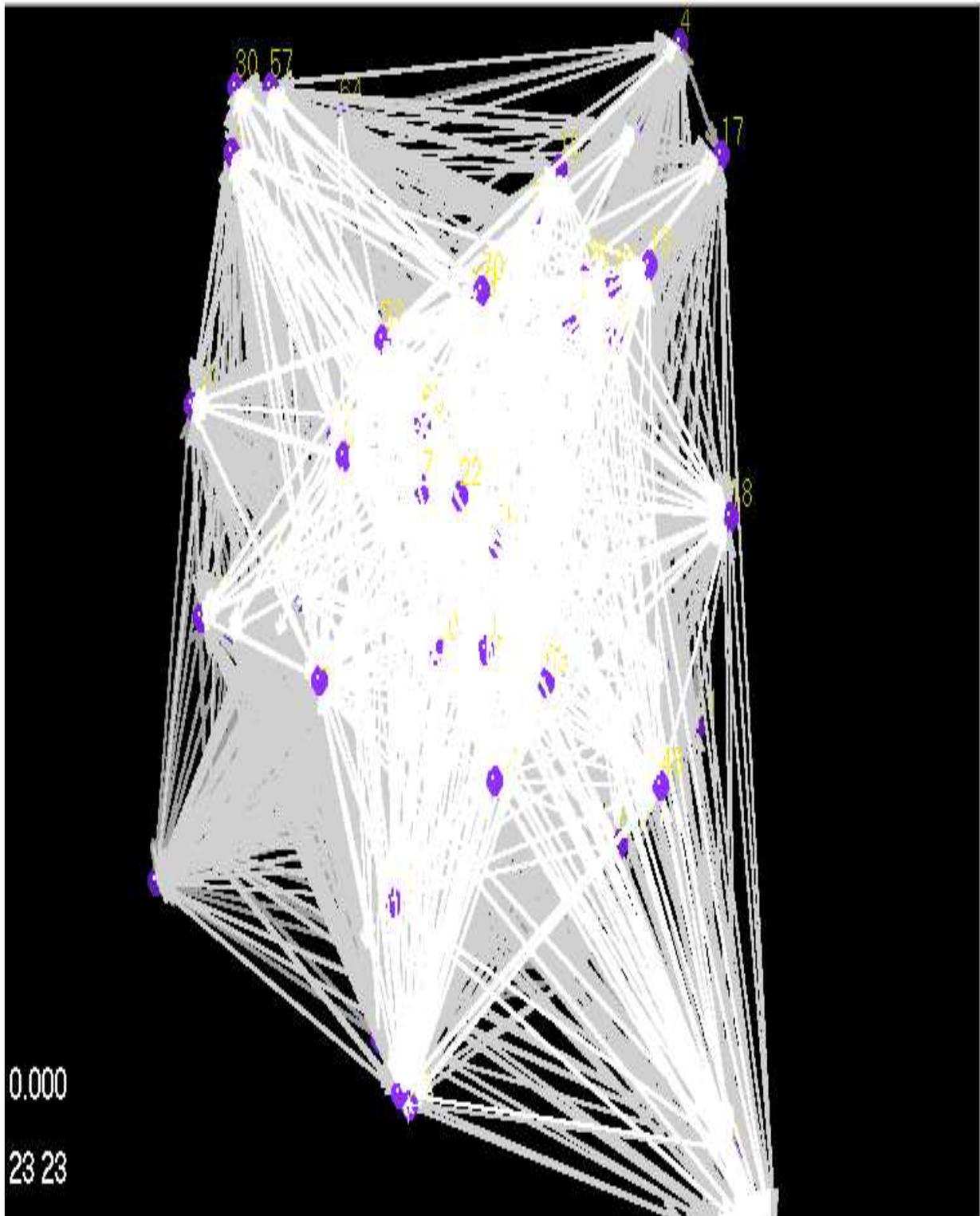


Figure 2- Interlinking within the research network



Figure 3- Examples of Anchorage in the research network

The research network—A linguistic analysis

Linguistic social networking theory seeks to describe the variations of a language in relation to their social network. Coates (1993:88) in her chapter on Social Networking in *Women, Men and Language* describes networks in highly industrialized, socially mobile societies as typically having low density and uniplex¹¹ relationships. The research network, as shown in figure 2, does not follow these typical patterns of social networking theory as prescribed by Coates. Conversely, this network appears to be a closed, dense network made up of multiplex¹² network ties. It has a clear dense core and a surrounding shell consisting of individuals who are related more uniplexly. The feature that makes blogging networks especially unique is the clearly defined outlet in the lower right hand corner. This outlet is the link to a second zone, or the area in which relationships are defined around another clusters of individuals. The first zone is a tightly connected group that may not know everyone in the zone, but theoretically could through linking between the zones. This second zone is nearly opposite to the first, in weblog communities. It represents a loosely structured network where links are often uniplex in nature. This zone, nevertheless, is also where electronic discourse occurs. For a more detailed description of zones, see below.

According to Coates, a closed network is formed by a group in which most of the contacts know each member of the network. This interrelation is true of bloggers, with the exception of the zone outlet. The outlet allows new bloggers to receive a reference, yet not necessarily become a core member of the group. One possible explanation for this outlet can be found within Lesley Milroy's theory of zones (Milroy 1987). Milroy's theory of zones takes into account two types of relationships, primary and secondary, labeled first order and second order. This theory is based on anchorage¹³ where each person is the focus from which relationships originate¹⁴. Milroy describes the zones as such:

“Each person may be viewed as a focus from which lines radiate to *points* (persons with whom he is in contact). These persons who are linked directly to ego may be characterized as belonging to his *first order* network zone. Each of these people may be in contact with others whom ego does not know, but could

¹¹ Relationships consisting of individuals only connected in one way.

¹² Relationships consisting of multiple types of ties.

¹³ The idea that a network can be considered from the point of view of individuals.

¹⁴ see figure 3.

come into contact with via his first order zone. These more distantly connected person's form ego's *second order zone*" (Milroy, L 1987:46).

Links in a blog occur in two places, the sidebar, also known as the blogroll, and the main post. The blogroll contains links of the first order zone in that bloggers only place the links of those that they read often or find interesting. Within the blogroll, *incestuous*¹⁵ linking occurs, creating a dense network of bloggers that read each other regularly. Within the posting area, linking of both first and second order can occur. This is the area in which bloggers link to information, or to blogs that post about a topic of interest, irregardless of placement in the network. It is also in this area where a relatively unknown blogger can enter the close network. The link in the posting area works like a *friend of a friend* tie in social networking theory. Milroy sites the importance of this role in gaining access to a network as, "first to guarantee good faith, but also as a means of becoming rapidly enmeshed in the rights and obligations relationships of the network" (Milroy 1987:54). Second order network links can only occur in the posting area. These links are by nature more impersonal than the blogroll and would not meet the qualifications of the dense, multiplex network represented in the blogroll.

Interesting assumptions can be made about the language of the research community by the type of network they represent. Dense, multiplex networks have the ability to enforce linguistic norms. These norms spread through hyperlinks with the same effect an echo chamber produces. In order to support the assumption that this network is closed and dense, two different features typical of blogging were examined; the hyperlink and the use of in-group/out-group language.

The hyperlink

To examine the hyperlink's role in blogging language, 10 blogs were randomly chosen from the research network and their links examined. Two different classifications of links were found; links of solidarity and informational links. Links of solidarity include references based on first name or references within the group. For example, if a name that is found on the blogroll, a member of the first order zone, is referenced by first name and hyperlinked, that would be considered a link of solidarity. Examples of solidarity links could often be found in the comment area. These links were often contained phrases of support and/or argument and were posted by the linked rather than the author of the blog the link if posted to.

An example of a link of solidarity:

”You May Be An Information Architect If...[Jeff](#) and I are in San Antonio, TX, on a consulting gig. Yesterday after work, we went out for beers with the team we’re working with. Jeff and I, being Bay Area liberals...”(excerpt from Peterme.com, April 22, 2003)

Links of information are references to articles, information, or people found outside of the network. If a link of information referenced a person, both the first and last names were given in the link.

An example of a link of information:

”[Michael O'Connor Clarke](#), Ireland's gift to Canada, writes:

[Could P2P music sharing actually be considered legal in Canada?](#) This tech journalist thinks so, and he makes an interesting argument. He's not a lawyer, of course - but it's an entertaining thought (excerpt from Joho.com, September 20, 2003).

Due to the fact that news and conversation spreads through the blogosphere most easily through links of solidarity, these types of links were expected to be prominent. The rate of solidarity versus the rate of information depended on the type of social networking functionality available. When building a weblog, one must take into account how social interaction should or should not be integrated into one's personal publishing space. If the author does not wish to receive feedback, he or she would chose software that does not support social features, or hand code his or her blog. Blogs that do wish to receive feedback from the audience embed social software into their weblogs. There are several ways to do this. One can add discussion forums, posting boards, or a guest book. By far the most common way to receive immediate and frequent feedback is through the use of comment and trackback features. Comment software allows the reader to make a comment on each topic individually. Many programs have begun to embed this feature in their software. Those that do not, often have open template code for plug-ins. Trackback, on the other hand, is a special feature of the program, Movable Type which records the web address of the blogs that have linked to your blog. Aptly entitled, *continuing the discussion*, trackbacks allow the blogger to keep a record of the on-going discussion his or her post may have sparked.

¹⁵ ¹⁵ “The *unbearable incestuousness of blogging*: “The other people who have blogs...read your blog, and it they like it they blog your blog on their own blog” (Clark, 59) (We’ve).

The blogs that used comment functions had a very high rate of solidarity. Blogs that did not allow for this had very high rates of informational hyperlinks. Of the 10 blogs randomly sampled, 7 used commenting features. Blog 1 and 2 use software called Movable Type which has social networking functionality embedded in the program. These functions allow other bloggers to comment and send messages that let the program know someone has just linked to them. Blog 1 had 80% links of solidarity compared with 20% links of information. Blog 2's rate of solidarity versus information is comparable to those without embedded social networking capabilities. This can most likely be attributed to blog 2's placement in the social network. The closer to the periphery of the social network, the lower the number of solidarity links.

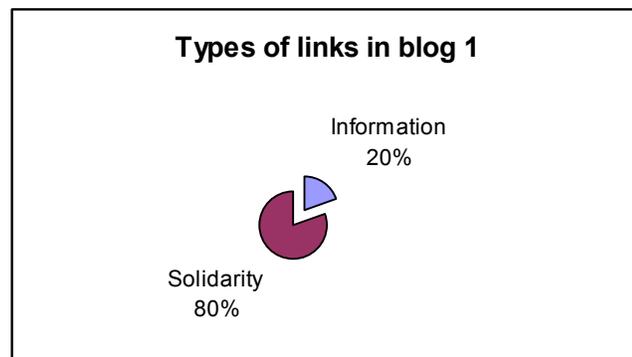


Figure 4-The types of links found in Blog 1

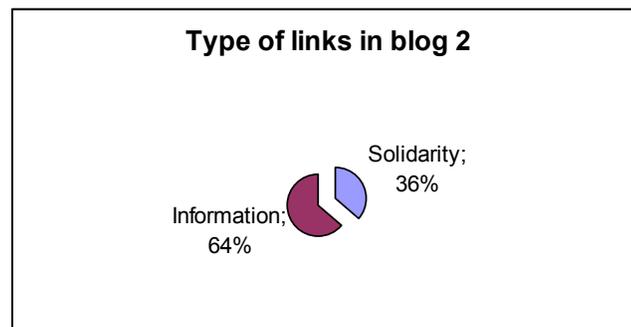


Figure 5- The types of links found in blog 2

Conversely, blogs 3 and 4 are created with blogger.com. Blogger does not have embedded social networking utilities, but free plug-ins are available in HTML format that allow the blogger to use comment functions. These blogs, however, has chosen to forego this feature. Their rate of solidarity versus information is very low. In blog 3, 64% of the links found were informational. Only 36% of all hyperlinks could be considered links of solidarity.

Similarly, 75% of all links counted in blog 4 were links of information; the remaining 25% were links of solidarity.

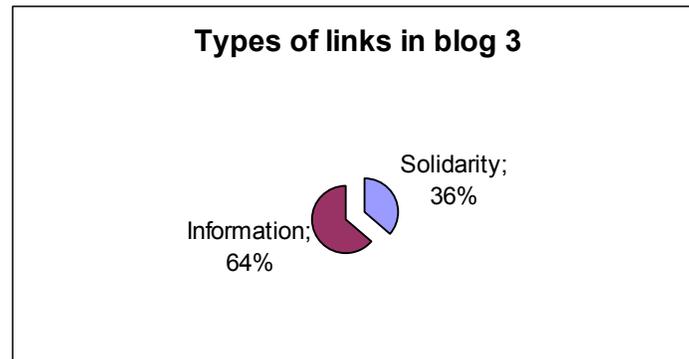


Figure 6-Types of links found in blog 3

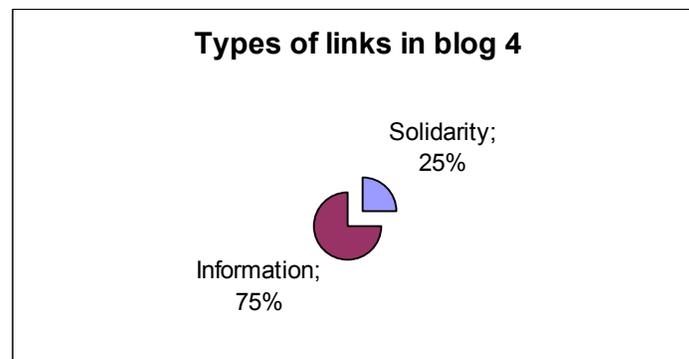


Figure 7- The types of links found in blog 4.

The type of link used plays a role in how the blog is placed within the social network. Blogs with a high rate of solidarity tend to be located in the center of the network, while those who use more informational links tend to be found on the periphery of the group. For example, blog 1 had a 17% higher rate of reciprocal linking (see figure 2 for a representation of reciprocal linking) from its peers than that of blog 4.

In-Group/Out-group language

In a closed, dense network, in-group language is often a prominent feature. In-group language reflects the unity within the group and serves to strengthen network ties. Blogs, seem to be an exception. Blogs are written in the first-person and often use links rather than in-group language to signify solidarity within the network. In this specific research network, I used Wordsmith to count the number of occurrences of certain keywords. In-group language keywords consisted of the personal pronouns, *we*, *us*, and *our*. Out-group language keywords consisted of the personal pronouns, *they* and *them*. As is evident from the table below, the rate

of out-group language is significantly higher than in-group language. The position of a blogger within the social network does not seem to play a large part in the usage of in-group and out-group language. Similarly, the length of a post does not seem to be a contributing factor. Extended posts did not show a significantly higher percentage of in-group language usage than posts of shorter length. A discrepancy between the typical linguistic attributes of the type of network graphically represented, closed and dense, and the type of language being used in this particular network was beginning to emerge. The research network was not adhering to the prototypical language features of a closed network, but using language in a new and exciting way.

occurrences of...	Blog 1	Blog 2	Blog 3	Blog 4
We	0.32%	0.57%	0.04%	0.43%
us	0.13%	0.18%	0%	0.08%
our	0.09%	0.15%	0%	0.15%
<i>TOTAL IN-GROUP</i>	0.55%	0.90%	0.04%	0.65%
They	0.53%	0.23%	0.47%	0.59%
Them	0.12%	0.13%	0.11%	0.23%
<i>TOTAL OUT-GROUP</i>	0.65%	0.36%	0.58%	0.82%

Figure 8—rate of in-group/out-group references in two blogs.

total in-group references: 0.54%

total out-group references: 0.60%

Part IV—The Post

The Basic Post—An analysis of basic linguistic attributes found in the main posting area.

Weblogs are nearly always written in the first person, and fairly informally. Whether written by a professional writer in his or her off-time, or by an online enthusiast, weblogs maintain a simple construction.

- “The local farm store had a deal on baby artichokes. So we took home a dozen, chopped some shallots, and then added the artichokes, salt, pepper, thyme, 3/4c white wine, a little water. Let it sit in the oven for an hour or so. Eat some now, save some for later. Yumm “ (Excerpt from –Insert Mark’s web address, April 29, 2003).
- “Another wedding invitation: two names and a date. A tiny red confetti heart held safe inside laminated plastic that will last forever” (excerpt from Jill/txt., April 29, 2003).

- “Yon Adaptive Path way, I've just published "[Finding the Right Users](#)", a guide to getting the right people to participate in your user research. While I wasn't blogging, I published "[Conducting International Usability](#)", my baldfaced attempt at getting someone else to fly me around the world to conduct user tests. And hey, super-catchy essay titles, eh? I suppose they make it clear what to expect” (excerpt from Peterme.com, April 3, 2003).
- **More goodness**
Via Tosca (who ordered it), this book made its way to my desk today: [Developing Online Games: An Insiders Guide](#)
- even more interesting designer stuff to lift from it - and nice tables like Jonathan Baron (X-box online producer) comparing single-user games with multi-player games.
Oh, and the book is written by two women: Jessica Mulligan and Bridgette Patrovsky” (excerpt from Klastруп’s cataclysms, April 3, 2003).

As the four excerpts¹⁶ above show, blog posts are usually written in short, paratactic sentences. They employ informal, non-standard constructions and slang such as “and hey, super-catchy essay titles, eh?” and “-even more interesting designer stuff to lift from it”. Both blogging jargon and professional jargon are also present in these excerpts such as *blogging*, *user tests*, *single-user games*, etc. Post lengths run between one or two words to hundreds of words. Length not only between weblogs, but also between posts varies greatly.

Due to the number of writers, the publicness of their work, the sheer amount of writing they produce, and the medium in which bloggers are writing, a jargon specific to blogging is evolving. Words like *blog*, *blogger*, *blogging*, *memes*¹⁷, *post*, *posting*, *trackback*, etc¹⁸ have become common place. In a closer look at the 10 most commonly used nouns in the research network, a significant percentage of these words are blogging jargon*. The remaining nouns often reflect the jargon of the researchers’ field of study**. The use of jargon in this particular research group appears to reflect group identity, both as a blogger and as an academic.

¹⁶ These four weblogs are not the same four weblogs mentioned during closer linguistic analysis.

¹⁷ Ideas and information with which ideas spread from blog to blog. (Mortensen & Walker: Online publication, no site given:262).

¹⁸ For a full list of blogging jargon, see <http://www.samizdata.net/blog/glossary.html>

	<i>Blog 1</i>	<i>Blog 2</i>	<i>Blog 3</i>	<i>Blog 4</i>
Most	Blog*	Tinderbox*	Time	Students**
Commonly	People	Hypertext*	Multiplayer**	Blog*
Used	Course**	People	Role**	Information
	Post*	Time	Worlds**	People
	Students**	Weblog*	Thesis**	Something
	War	Work	Conference	Time
	Blogger*	Weblogs*	Apprentice**	World**
Least	Blogs*	Summit**	Book	Course**
Commonly	Weblog*	World	English	Blogs*
Used	Friends	Ideas	Weblogs*	Thesis**

Another linguistic feature of blogging posts that became apparent after the creation of word lists from the research network corpus was the use of frames. “At a very general level, the notion of a ‘frame’ provides an attractive metaphor for thinking about discourse understanding as, at least partially, ‘a process of fitting what one is told into the framework established by what one already knows’...the basic structure of a frame contains labeled slots which can be filled with expressions” (Brown & Yule, 1983:239). For example, this excerpt from Jill/txt, “Today I scored my first free pizza in compensation for technical assistance. Peppe's Pizza very wisely provides wireless, thus making them the ideal place for this kind of transaction. Do I get to call myself a geek now? Or at least the MoveableType guru of Bergen?”(April 16,2003). In order for the reader to fully understand this post he or she must have the frame for installing computer programs, possible with a slot for the specific program mentioned, Movable Type. Those outside of the blogging community would not be likely to understand the difficulties often occurring when trying to initially install this specific program. Those who do have that background knowledge understand why pizza was an appropriate compensation for Jill’s time and labor. Framing allows the author to make the assumption that his or her readers will have the background knowledge for the presented concept so that he or she does not need to explain it, again lessening the need for many adjectives.

PART V—Blogspeak

Looking at blogs from a completely sociolinguistic standpoint was leading to more questions than answers. There is a clearly defined social structure in weblogging communities, a defined sense of community, and an obvious on-going conversation; so how can one account for the variety of language being found? Weblog variety was using a mixture of socially strengthening language with language that is often found in loose, uniplex networks or even in *non-network* communication. The answer was suggested by the function of the weblog. Blogs combine both the monologue and the dialogue in a space-bound, electronic environment. They are simultaneously self-reflective thoughts presented publicly and continuous conversations. Blogs utilise both the attributes of on-line, informal *spoken* language with those of the conventional written monologue. Again, it should be stated that this study was conducted on a small scale. Further research is necessary to corroborate conclusions as to the attributes of blogspeak.

Before delving into the type of language specifically used in blogs, it is necessary to take a look at the variety of language commonly associated with the Internet, *netspeak*¹⁹. While blogs are clearly a written medium, they also have spoken qualities. For instance, weblogs are socially interactive, immediately revisable, and somewhat spontaneous. Many prominent bloggers consider blogging a form of writing with peppered with frequent bursts of conversation. Examining *netspeak* is necessary in order to understand how it and other conventional forms of communication have merged to create this new variety found in blogs.

Netspeak

As this is not a study to differentiate or define speaking and writing, but rather to identify the general linguistic characteristics of a blog and the sociolinguistic factors that play a role in that language, I am using David Crystal's definition as charted in table 2.1 of his book, *Language and the Internet*.

¹⁹ *Netspeak*, according to David Crystal is “a type of language displaying features that are unique to the Internet... arising out of its character as a medium which is electronic, global, and interactive” (Crystal, 18).

Differences between speech and writing (from Crystal, 2001)

Speech	Writing
<ul style="list-style-type: none"> ✓ Speech is time-bound, dynamic, and transient. ✓ There is no time-lag between production and reception, unless one is deliberately introduced by the recipient. The spontaneity and speed of most speech exchanges make it difficult to engage in complex advance planning. ✓ Because participants are typically in face-to-face interaction, they can rely on extralinguistic cues. ✓ Many words and constructions are characteristic of (especially informal) speech, such as contracted forms (<i>isn't, he's</i>). There is nonsense vocabulary (e.g. <i>thingamajig</i>), obscenity, and slang. ✓ Speech is very suited to social or 'phatic' functions. It is also good at expressing social relationships, and personal opinions and attitudes, due to the vast range of nuances. ✓ There is an opportunity to rethink an utterance while the other person is listening. Interruptions and overlapping speech are normal and highly audible. ✓ Unique features of speech include most of the prosody. The many nuances of intonation cannot be written down with much efficiency. 	<ul style="list-style-type: none"> ✓ Writing is space bound, static, permanent. ✓ There is always a time-lag between production and reception. Writing allows repeated reading and close analysis, and promotes the development of careful organization and compact expression, with often intricate sentence structure ✓ Lack of visual contact means that participants cannot rely on context to make their meaning clear; nor is there any immediate feedback. ✓ Some words and constructions are characteristic of writing. Certain items of vocabulary are never spoken, such as the longer names of chemical compounds. ✓ Writing is very suited to the recording of facts and the communication of ideas, and to tasks of memory and learning. ✓ Errors and other perceived inadequacies in our writing can be eliminated in later drafts without the reader ever knowing they were there. ✓ Unique features of writing include pages, lines, capitalization, spatial organization, and several aspects of punctuation

Figure 9- (cont.) Differences between speech and writing (Crystal , 26-8).

What then is Netspeak?

Netspeak, the language of the Internet, faces many of the same linguistic problems as does language in the ‘real world’. Netspeak can be seen in a unique way, however. If one considers all the different varieties of written and spoken communication to be the parent language of the human race, one can look at the language of the Internet as the child of this union. This language took on the characteristics of the human mother tongue, written and spoken forms of communication, and assimilated it into the constraints of a digital environment²⁰. Netspeak relies on both the criteria found in speech and writing. Many style guides, like *Wired Style*, suggests writing as one speaks. One must keep in mind, however, that the Internet is a global community. The world is made up of an infinite number of dialects, situated upon a *dialect continuum*²¹. As the Internet mirrors the real-world environment, it also contains the same infinite number of possibilities for social dialects.

If we take the above criteria and compare it with Netspeak, in particular the language of blogs, we can develop a working definition on which to base a linguistic study of blogs. I will use the term *blogspeak* to define the variety of language which is native to the unique format found in weblogs.

Spoken Language Criteria	The Web (in general)	Blogs (in general)
1. Time-Bound	No	Yes
2. Spontaneous	No	No
3. Face-to-face	No	No
4. Loosely Structured	Variable	Variable
5. Socially Interactive	No, with increasing options	Variable
6. Immediately Revisable	variable; depends on the tools.	Yes
7. Prosodically Rich	Variable	Variable, but limited

Figure 10- Adapted from Crystal's Spoken Language Criteria applied to Netspeak (Crystal, 42)

Written Language Criteria	The Web (in general)	Blogs (in general)
1. Space-Bound	Yes, with extra options	Yes
2. Contrived	Yes	Yes
3. Visually Decontextualized	Yes, but with considerable adaptation	Variable

²⁰ For a discussion on these constraints see Crystal's chapter on Netspeak in *Language and the Internet*, 2002.

²¹ A dialect continuum is referred to as a series of different dialects which gradually merge into one another (Trudgill, 30).

4. Elaborately Structured	Yes	Yes
5. Factually Communicative	Yes	Yes
6. Repeatedly Revisable	Yes	Yes
7. Graphically Rich	Yes, but in different ways	Variable, but limited

Figure 11- Adapted from Crystal's *Written Language Criteria* applied to Netspeak (Crystal, 43)

These tables attempt to show that blogs combine features of both spoken and written communication to create a unique format. Seemingly contradictory terms are not contradictory in blogs. For example, while blogs are loosely constructed in that they are informal, they can also be elaborately structured in that blog posts often support threads of conversation as well as internal monologues. More detailed descriptions of each criteria can be found below.

If we continue with the metaphor of speech and writing as the ‘parents’ of netspeak in general, and blogspeak in particular, we could draw on the scientific theory of DNA replication²² to explain the linguistic union of these two parent forms of communication. Speech and writing are both seen as parent strains of the Internet’s *lingua franca*, Netspeak. When these parent strains unite, a new strain of language is produced. This new strain takes qualities from both the parent strains and applies them as necessary in its new environment. Blogspeak, however, is somewhat different than Netspeak in that it combines writing and speech in a unique environment, one that supports both the written internal monologue and threads of conversation. Netspeak, on the other hand, must be more adaptable as the Internet is comprised of many different genres of language, each uniquely adapted for its environment, similar to the theory of natural selection.

Blogspeak—the language of the weblog.

Every environment within the Internet has its own format, and within each format evolves a specific dialect, tailored to the abilities and limitations native to that space. Blogs are generally accepted as having a form over function relationship in that a blog is often defined singularly by its form. For example, a research blog functions as an online record of research, as well as a medium to share and informally publish information with ones colleagues. A personal weblog log that describes the daily life of an individual is quite different functionally, but both are considered blogs of equal status due to their format.

²² For more information about gene inheritance, see *DNA Replication*, <http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/D/DNAReplication.html>

Within this form, language has evolved to accept standards of writing, as well as the standards of speech to form an evolving variety of language, unique to the online environment.

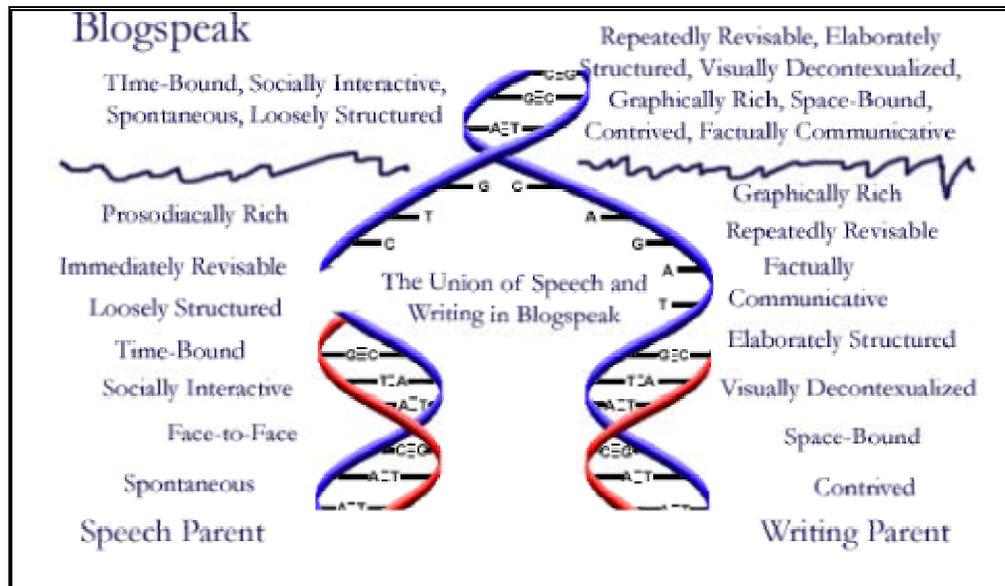


Figure 12- The Union of Speech and Writing in Blogspeak

The picture above is based on the theories of speech and writing by David Crystal as stated in the previous section. Blogs acquire these attributes because of the nature of the format of the medium. Blogs are short entries, frequently updated, typed into a computer, and bound by the restraints of the keyboard or program one uses to record the blog.

The new linguistic attributes of blogging are:

From the speech parent:

1. Time-bound: The language of blogs is time-bound in that they are limited to a certain accepted length of post. They are only presented on the front page of a blog for a certain period of time before they are removed from view to the archives. Blog posts, however, are still available to those who search the archives. In that respect, blogs are not as time-bound as speech, yet more time-bound than non-archiving, static webpages.

2. Contains prosody: Ability to express nuances of speech, as expressed by illusions of speech (see below) through the use of non-standard punctuation and emoticons. The abilities to use avi and midi files within blog posts also contribute to the ability to contrast loudness, rhythm and tempo. While avi files and midi files are not the main ways of communicating in a blog, more and more bloggers are using these capabilities to present themselves in artistic ways that were previously impossible with strictly text-based weblogs. For example, poetry can be

given soundtracks and paintings can be enhanced by music. Weblogs are also using orthographic euphemisms in order to disguise language from search engines. For example, different spellings of *porn*, (pOrn, prOn, etc) are used in order to keep one's blog from being blocked by search engines with child safe functionality. Using blogs as a multimedia platform is still in its infancy, however. Blogs remain very textual, yet have the capabilities to prosodically rich.

3. Socially Interactive: Blogs are a form of social software. They are the medium by which threads of conversation spread through the blogosphere²³.

4. Spontaneous: A feature of blogging is its lack of editing. Most posts are quickly jotted down notes intermingled with frequent bursts of conversation.

5. Loosely structured: Blogs are generally accepted as a place to think things through, a place to organize thoughts. This organization happens through several posts and through the feedback of other bloggers. Concrete structure is not prominent in each post. Jill Walker of, *jill/txt* says this about the structure of blogs: “research happens in blogs, and in the conversations between blogs. Blogs aren't about documentation, they're about doing, thinking and discussing. And they're about catching fleeting thoughts and making them explicit” (http://huminf.uib.no/~jill/archives/blog_theorising/not_documenting_doing.html).

From the writing parent:

1. Repeatedly revisable: Blogs are unique in that the author can edit as often as need be. One catch in the ability to revise blogs is the google cache. Google, the popular search engine, stores information on everything they list. Even if a blogger deletes an entry, it is still possible to find the entry on google.
2. Elaborately structured: Blogs combine the attributes of the internal monologue, the dialogue, speech and writing. They foster threads of conversation through linking and support a rich social network. Blogs do this by using language to its fullest electronic sense. Language found in blogs can contain several subordinate clauses, or read like a short staccato conversation. It can use nonsense words like *whachmacallit*, abbreviations

²³ “Blogosphere” is a term used by bloggers to describe the entire network of blogs; the ecosystem of blogs.

like LOL (laughing out loud), or spell out the nearly unpronounceable genus name of the Dandy Lion.

3. Visually decontextualized: with the exception of the rarely used video file, blogs do not provide paralinguistic cues.
4. Graphically rich: blogs have the capability to be graphically rich in that many bloggers use pictures in their main content area, as well as maintain photoblogs²⁴. Hyperlinks also provide a way to introduce graphics through a secondary source.
5. Space-bound: Although vast, space on the web is not infinite. Server restraints and upload times limit the full range of linguistic possibilities. As servers become faster and more room available at affordable prices, and as multimedia equipment such as mobile blogging and camera equips cellular phones continue to decrease in price and increase in quality, more multimedia files such as midi and avi files may change the discourse of blogging. Blogs are also space-bound in that they are limited to the display screen. Just as text is limited to the space in which it is displayed, blog text is limited to the screen on which it is read.
6. Factually communicative: blogs communicate facts through discussion and through hyperlinking. A fact can be thoroughly discussed to the extent of the bloggers knowledge, then hyperlinks can be added to reinforce or add information.
7. Contrived: Blogs are contrived! There is a time lag between the posting of a blog and the response, if any is received. The internal monologue of a blog anticipates this by allowing hyperlinks to develop his or her argument. A dialogue post must anticipate this need as well, often using the hyperlink in the same fashion. Comment functions allow the blogger to respond to confusion, but the lag time between the production and reception of the post may lead to the necessity of several posts to alleviate general confusion.

²⁴ A photoblog is a weblog only consisting of pictures and captions.

Oral and literate strategies:

Borrowing from both the parent strains of speech and writing, blogs form a type of communication that takes into account the abilities and limitations of the internet to create a variety that will harness the communicative abilities native to on-line discourse.

Far from being limited to the above features, blogs contain the characteristics of both oral and literate *strategies*. These strategies have combined in blogs where both the monologue and the dialogue create a distinctive atmosphere of conversation and internal self-reflection. Deborah Tannen, as discussed in John S. Schafer's paper, *Linguistic description of speaking and writing and their impact on composition pedagogy*,

defines oral strategies as “aspects of discourse which make maximal use of context, by which maximal meaning and connective tissue are implied rather than stated”, written strategies, on the other hand, are aspects of discourse “by which maximal background information and connective tissue are made explicit...oral strategies...depend for effect on paralinguistic and nonverbal channels, while literate strategies...depend on lexicalization to establish cohesion.”(Schafer 1983:10).

While the majority of blogs are a written medium, they often combine oral and literary features to make meaning explicit. One such feature, arguably the most prominent meaning making feature in a blog is the link. Links make meaning, beyond the function of network identification, explicit through their scripting function. A script, or frame, is the “cognitive structure...knowledge of which is presupposed for the concepts encoded by the words” (Ungerer, Schmid 1996: 209). The link is used to make references that are otherwise implicit, explicit. For example, if a person links to an article on CSS stylesheets, the link takes away the need to explain what a CSS stylesheet is because the example or an explanation is just a click away. This leaves the blogger to post about the reason he or she is linking to this page, inferring that the reader either has, or can acquire required background knowledge necessary to understand the communication. The links create a cognitive chain that come together to create a whole picture, or a script.

An example of such usage can be described as such:

1. Excerpt from blog, [steffanie.net](#):

“I posted some photos from the Blog Dinner in Malmö in my photolog [starting here](#). Gustav posted a [summary](#) of the evening's discussions and [a lot of photos](#). Mats posted not [one](#) or [two times](#), but [three times](#) about the dinner!” (<http://www.steffanie.net/blog/arkiv/000502.html>).

2. Possible reader questions:

- a. What is a blogtalk dinner?
- b. Who is Gustav?
- c. Who is Mats? Why is he posting so much? What did he say?

These questions are answered through the use of linking. It is not necessary for Steffanie to be explicit in this post because all the background information that the reader may or may not have can be obtained through one click to the right permalink²⁵.

Links also help in the combination of internal monologue and dialogue. An internal monologue may consist of something that is private, intended to help clear one's thoughts, or a random reflection without an intended audience. By adding links to the monologue, the blogger is taking this internal monologue and outfitting it with the information the reader needs to internalize the post and reply. The link has become an invitation to conversation as well as clearing up common writing-related ambiguities with context. These types of posts are generally replied to from those within a bloggers first order zone, the majority being reciprocal links of solidarity.

Blog writers often converse using the *illusion of speech* (Schafer 1983:18), as described by Page in Schafer's paper, *Linguistic description of speaking and writing and their impact on composition pedagogy*, rather than transcribing actual speech.

“Intellectually turned on...

Few things in life give me greater joy than witnessing (if only from afar) another's intellectual excitement and imagination. [Matt Webb](#) is currently reading D&G's [A Thousand Plateaus](#) and taking good notes.... But beware Matt - once you start quoting D&G people treat you differently ;)” (excerpt from http://www.purselipsquarejaw.org/2003_06_01_blogger_archives.php#200402469).

This excerpt from Anne Galloway's page, Purse Lips, Square Jaw, is an example of indirect dialogue. It contains other methods of speech presentation, as suggested by Page:

1. *Indirect, submerged speech*: Anne is indirectly complementing Mat on his reading choice and intellectual ability.
2. *“Slipping” from indirect into direct speech*: Direct speech to Mat

²⁵ A permalink is a link to a posts permanent location in the archives.

3. *Parallel indirect speech*: Anne is witnessing intellectual excitement from Mat's reading choice.
4. "*Colored*" *indirect speech*: The title suggests sexual excitement from Mat's reading choice as well as from the opportunity to witness it.

The use of other items such as emoticons, italics, lexical items, capitalisation, fragments, and dialectal and colloquial forms can all be found to create the illusion of speech in blogs. The speech may be an illusion created by bloggers, but the continuous conversation is real.

Conversation, or dialogue, is prevalent in weblogs. Weblogs converse through several means. The most obvious way to converse is through comment features and trackbacks. Comment features are attached to every post individually and allow readers to comment on a specific topic, rather than the page as a whole²⁶. This type of commenting is often reciprocal, the author of the weblog responding to comments via a new post, or in the comment section itself. Less obvious conversations, and harder to track, exist in the main posting areas. Posts that reference another blogger and/or comment on his or her post create threads of conversation. These threads are picked up by others in the social network and commented on in their blog. This is where software driven trackback functionality becomes important. Trackbacks are automatically detected by the computer and allow the blogger to be a passive participant in locating threads of conversation. Trackbacks allow the blogger to read the response of another author that has linked to them in their main posting area. This gives the blogger the opportunity to respond to the post and continue the conversation. Blood has likened this process in a keynote speech at BlogTalk, Vienna, 2003 to an echo chamber, and warns that this effect often limits participation from outside the social network and does not allow for freedom of expression from those that are not *native* to the network (Blood, <http://blogtalk.net/presentations/wagingpeace.pdf>).

Another example that conversation does in fact exist in weblogs is the presence of H.P. Grice's Maxims of Conversation. These maxims are not only explicit, but also encroached upon for effect as they would be in a real-world situation. For example, the maxim of quantity is observed through the use of hyperlinks. The blogger posts all that he or she needs to say, then gives the hyperlinks for any additional information the reader may or may not need. This maxim is strengthened by the function of many blogging tools. Most tools give a box in which to write your post. These boxes are not large, but they are expandable. Bloggers also reinforce this maxim by writing short posts often. Blogs are not

²⁶ Commenting on an entire page would be done through a guestbook.

by nature long-winded. Like conversation, blogs do not discuss many different topics in a short span. Conversation will change topics, but with obvious markers and compliance on the part of all participants. Blogs too write short posts under topic headings. When the topic changes, a new post, or topic heading is made. The topic can be encroached upon for effect, as in spoken conversation. A blogger can rant and rave about how angry he or she is, going on and on until the maxim of quantity is destroyed.

Conclusion

Within the blogosphere, a certain variety of language has developed, one that rationalizes the human need to communicate with the constraints of the medium. These weblogs group into tight social networks that consist of varying degrees of community. Unlike traditional dense social networks, weblog communities do have an outlet, a portal, which allows for the entrance or exit of community members. Each community is anchored around a topic, whether it is job-related, research-related, age, gender, etc. Because of the closeness of the blogging networks, varieties of language can be standardized, strengthening further the sense of group identity.

Discourse devices like the monologue and the dialogue come together to create extensive conversations that spread through the blogosphere with an effect not unlike an echo chamber. These threads of conversation use many of the linguistic features found in both oral and written forms of communication. The two major forms of communication have united to produce a variety that utilizes the full capabilities of the Internet environment as a means to converse and reflect. Because of the constraints of length and the purpose of this essay, this study was approached on a small scale and only within one research community. Further research is needed in order to document the level of standardization within weblogging communities, and the blogosphere as a whole.

Bibliography:

- Barab, S.A., Kling, R., & Gray, J.H. (Eds.). (in press). "Designing for Virtual Communities in the Service of Learning." New York: Cambridge UP.
- Blood, Rebecca. The Weblog Handbook: Practical Advice on Creating and Maintaining Your Blog. Perseus Publishing, 2002.
- Blood, Rebecca. "Weblogs: A history and Perspective." We've Got Blog. Comp. and ed. John Rodzvilla. Cambridge: Perseus Publishing, 2002.
- Brown, Gillian, and George Yule. Discourse Analysis. Cambridge: Cambridge UP, 1985.
- Bucholtz, Mary, A.C. Liang and Laurel A. Sutton (Eds). Reinventing Identities: The Gendered Self in Discourse. Oxford: Oxford UP, 1999.
- Clark, Joe. "Deconstructing 'You've Got Blog'." We've Got Blog. Comp. and ed. John Rodzvilla. Cambridge: Perseus Publishing, 2002.
- Coates, Jennifer. Women, Men and Language. 2nd ed. Harlow: Longman Group Limited, 1993.
- Crystal, David. Language and the Internet. Cambridge: Cambridge University Press, 2001.
- Davis, Boyd H., and Jeutonne P. Brewer. Electronic Discourse: Linguistic Individuals in Virtual Space. New York: State University of New York Press, 1997.
- Dibbell, Jullian. "Portrait of the Blogger as a Young Man." We've Got Blog. Comp. and ed. John Rodzvilla. Cambridge: Perseus Publishing, 2002.
- Frankel, Mark S., and Sanyin Siang. Ethical and Legal Aspects of Human Subjects Research on the Internet. A Report of a Workshop, June 10-11, 1999.
<<http://www.aaas.org/spp/dspp/sfrr/projects/intres/main.htm>> New York: 1999.
- Graham, Brad L., "Why I Weblog: A Ruminaton on Where the Hell I'm Going with This Website." We've Got Blog. Comp. and ed. John Rodzvilla. Cambridge: Perseus Publishing, 2002
- Katz, Jon. "Here come the Weblogs." We've Got Blog. Comp. and ed. John Rodzvilla. Cambridge: Perseus Publishing, 2002.
- Milroy, Lesley. Language and Social Networks. 2nd ed. Oxford: Basil Blackwell Ltd, 1987.
- Mortensen, Torill. "Re: your thoughts." E-mail to Stephanie Nilsson. 3 April, 2003.
- O'Grady, William, Michael Dobrovolsky, and Francis Katamba, (ed). Contemporary Linguistics. 3rd ed. Harlow: Addison Wesley Longman Limited, 1996.
- Schafer, John. S. "Linguistic Descriptions of Speaking and Writing and their Impact on

- Composition Pedagogy.” <<http://jac.gsu.edu/jac/4/Articles/7.htm>>. 1983.
- Taylor, John R. Linguistic Categorization: Prototypes in Linguistic Theory. 2nd ed. Oxford: Oxford UP, 1995.
- Trudgill, Peter. Sociolinguistics: An Introduction to Language and Society. 4th ed. Penguin: USA, 2001.
- Ungerer, F., and H.J. Schmid. An Introduction to Cognitive Linguistics. Harlow: Pearson Education Limited, 1996.
- Warschauer, M. (in press). “Language, Identity, and the Internet.” In B. Kolko, L. Nakamura, & G. Rodman (Eds.), Race in Cyberspace. New York: Routledge.

Websites Cited:

- Anne Galloway <http://www.purselipssquarejaw.org>
- Blogger: <http://www.blogger.com>
- Blogstreet <http://www.blogstreet.com>
- Blogging Jargon <http://www.samizdata.net/blog/glossary.html>
- DNA Replication
<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/D/DNAReplication.html>
- Jill/txt <http://huminf.uib.no/~jill/>
- Joho.com <http://www.joho.com>
- Movable Type: <http://www.movabletype.org>
- OED On-line: <http://www.oed.com>
- PeterMe: <http://www.peterme.com>
- Steffanie.net <http://www.steffanie.net>

Conference Cited:

BlogTalk - A European Conference On Weblogs: Web-based publishing, communication and collaboration tools for professional and private use. May 23-24, 2003. Vienna, Austria.
<<http://www.blogtalk.net/>>