

Identifying Linguistic Traits Between Internet Content Producers and Consumers

Jerrett Longworth

ENC 1102

Esteemed Professor Loren Cooper

October 21, 2019

### Identifying Linguistic Traits Between YouTube Content Producers and Consumers

YouTube is a prevalent medium for both producing and consuming content in a variety of topics. The freedom that YouTube provides as a platform allows anyone of any background to create content and share their passions. However, before understanding the discussion of contemporary internet discourse, a process of reasoning must be established from the universally familiar to the relatively less familiar. According to Deborah Brandt (2017), literary sponsorship is the idea that organizations can sponsor – or provide access to – literacy. Aside this idea of literary sponsorship, Brandt discusses that no one becomes literate on their own; rather, all people are dependant on literary sponsors to provide literacy. The idea of discourse communities lies the foundation for what influence organizations and corporations have on individuals.

Within the YouTube community, it could be said that content creators are in essence a literary sponsor. Audience members watch these videos and (intentionally or not) gain some type of literacy as a result. Whether it be a quite literal interpretation of this by learning a new language from YouTube or by picking up new colloquialisms from a specialized creator, audience members will take some of the content they see and reuse it in a different environment. For example, various content may offer true, legitimate information in explaining a topic, however could do so in a way that may be considered biased. Literary sponsorship, although important for discussion, is not the only lens through which to study this complex environment.

Ann Johns (2017) is another well-known linguist and discusses the idea of discourse communities. She defines the characteristics as essentially the following: a discourse community shares common goals, communicates through the same medium(s), uses the same literary genres,

and has a maximum threshold of membership. Each discourse community has their own sets of conventions and rules specific to that community. These conventions, in turn, help define the discourse community with identifiable traits. The discourse community of discussion is the YouTube discourse community – but more specifically – the discourse community that consists of the followers of Luke Smith, a creator of YouTube content.

Though it could be done with some assumption that the social atmosphere within the internet is already known to the reader, this would do a disservice to understanding the nature of a discourse community as complex and neoteric such as this one. To understand this nature, I propose to first understand the social climate of the internet in a more primitive state – one less influenced by other more recent outstanding issues such as internet censorship, trolling, and net neutrality.

Anandra Mitra and Eric Watts (2002) propose that one lens to view internet discourse is with a metaphorical lens of voice. Voice represents the individual, and with it comes the abstraction of voice onto a digital medium. The discourse space that exists within the internet (as Mitra and Watts refer to as cyberspace) has some unique properties that is unprecedented among any other form of written communication to date. As they put it, the internet has the “potential to flatten hierarchies of power,” wherein that the internet has no central location or time, as it contains all conversation from everywhere and all time. These special properties mean that credibility of internet discourse must be judged by audience members differently than if the same conversation occurred physically. Rather, internet speakers’ credibility is judged based on “emotional and ethical terms” (Mitra & Watts, 2002). Despite the outdated perspective, this serves as a fundamental line of reasoning to understanding contemporary internet discourse.

An example of modern internet discourse can be found in Daniel Lu's (2018) research of its social atmosphere from the perspective of a "standard user," which he does not hold formal certification or even extensive experience in rhetoric. Instead, Lu provides a perspective of the internet that is more native to the current social atmosphere. Essentially, users of one social media service, Reddit, have the ability to speak anonymously and without the pressure to uphold a public image like other social media services. With this in mind, attention can now be moved to other social media services, like the central one of discussion here: YouTube. YouTube's social atmosphere is incredibly complex due to its entanglement with many facets of modern society, so I believe taking incremental steps is crucial to understanding YouTube's discourse community. The collection of data and consequential analysis of YouTube's rhetorical connection between content creators and their audience members is the primary objective set forth here.

For supplemental context, Sabatini and Sarracino (2019) offer recent conjecture of possible effects of using social media. After their own analysis of surveys with 50,000 people, they determined that "the use of [social networking services] is negatively associated to people's propensity to trust strangers, neighbours, and the police." This means that people are less likely to trust others if they use social media more frequently. The study – while merely speculation – proposes that more research and analysis must be done of social media and modern internet discourse.

To succinctly overview the following research, I propose the following: What is the rhetorical connection between YouTube content creators and the audience members that consume said content? In attempting to understand this, I suggest that the underlying foundations

of modern internet discourse be established before progressing. Ann Johns (2017) and Deborah Brandt (2017) discuss literary sponsorship and discourse communities, respectively. These concepts are the overall foundation of understanding the context of YouTube's discourse. Audience members consume content produced by YouTube content creators, and therefore are in some extent being sponsored by YouTube. Additionally, YouTube audience members and content creators alike participate in specific discourse communities. Mitra and Watts (2002) build on this by transitioning from 'real' discourse to internet discourse. Since the internet fundamentally changes communication into a continuous conversation and the authenticity of users is no longer bounded by public image, this shifts credibility from rhetorical appeals to "emotional and ethical terms." A series of other more recent studies follow this idea as a way to capture surrounding details and effects of this ever-changing social climate.

## Annotated Bibliography

Ahmad, U., Zahid, A., Shoaib, M., & AlAmri, A. (2016). HarVis: An integrated social media content analysis framework for YouTube platform. *Information Systems*, 69, 25-39.

doi:10.1016/j.is.2016.10.004

HarVis outlines a method of “Harvesting and Visualization” based analysis for YouTube content. The writers of this paper, all of which are experienced programmers for institutions worldwide, created this software in an attempt for a more intricate system of collecting and analysing data from YouTube. The paper outlines the specific algorithms and processes that take place to collect and analyse data. The authors point out that many other current solutions for YouTube video analysis simply use the titles and descriptions of videos, but this does not encapsulate the full set of data available. HarVis also gathers data like view count, likes and dislikes, comments, connections between users, and timestamps. While this appears to be a solution that should be more widespread, there is very little documentation on this software and it appears that it requires knowledge about programming and compiling Java to use the software. Outside of a singular copy of the code, there is no detail on where this code came from or the development process behind it, which could decrease the legitimacy of it overall. However, this software is distributed in a way where all of the code is available without obfuscation, so any illegitimacy can be detected, should one read that code.

This paper is technical and, on its own, is not incredibly useful since the amount of time and labor required to construct software from scratch using this paper would be too great to be useful. However, the details in this paper have led to the code of the program itself.

In this way, it may be possible to compile and run the code to use for data analysis on YouTube videos. This could significantly expedite the time it takes to collect and analyse the primary source in this paper. With previous experience building programs from source code (albeit with documentation), this may prove to be a trivial task to apply the same knowledge to this instance. Otherwise, there are a few other methods of YouTube video data collection that I already understand and have used before, but they are not as comprehensive as software like HarVis. The authors itself recognise its potential to analyse “socio-linguistic features and other properties.”

Brandt, D. (2017). Sponsors of literacy. In E. Wardle, D. Downs, & J. E. Sullivan III (Ed.), *Writing about Writing: A College Reader* (pp. 68-81). Boston, New York: Bedford/St. Martin's.

In *Sponsors of Literacy*, Deborah Brandt discusses the idea of literary sponsorship, and that no one becomes literate on their own. Brandt is a professor in the Department of English at the University of Wisconsin, researching and writing books about literacy for at least 30 years. She describes “literary sponsors” as these entities that somehow allow others to obtain access to (or possibly be constrained from) literacy. The word “sponsors” is used because it both relates to the authoritative figures in people’s lives, and it connotes the commercial entities from which literacy can come from (radio, TV, etc.). Continuing forward, Brandt discusses a more relevant example of sponsorship and how it relates to access of literacy: school. Through a series of real-life examples of varying socioeconomic backgrounds, she shows the way that a school (or a job) and its own goals relate to access to literacy. If, for example, the need of a company changes to focus more

on slow, logistic processes to settle matters, then its employees will have to change their rhetoric to match the goals of the company in exchange for access to literacy in the first place.

In my own research, *Sponsors of Literacy* may show to be a helpful foundation for its perspective of the influence organizations and corporations have on individuals. Luke Smith, a creator of YouTube content, is in his own right a sponsor of literacy. Audience members watch his content and consequently may develop certain habits of literacy not unlike Smith's own. This is currently only speculation without the formal research conducted, but it would not be unreasonable to conject the possibility of audience members gaining literacy from YouTube creators. In that sense, it may be important to discuss the responsibility that these creators have as a result of their position online. For an elaborated example, while still conjecture, Felix Kjellberg (better known as PewDiePie) is the most-subscribed individual on YouTube. He exemplifies this responsibility to be mindful of the words and actions in his content. There are many impressionable people in the world that have very easy access to this type of content, and it could be found very quickly to be, to the possible dismay of parents globally, a source of new vocabulary and interests.

Johns, A. M. (2017). Discourse communities and communities of practice: Membership, conflict, and diversity. In E. Wardle, D. Downs, & J. E. Sullivan III (Ed.), *Writing about Writing: A College Reader* (pp. 319-326). Boston, New York: Bedford/St. Martin's.

In this piece, Ann Johns discusses the premise of discourse communities, their characteristics, and what implications this has for rhetorical analysis. Essentially, Johns



defines discourse communities of a group of members that share common goals, communicate through the same medium(s), and has a maximum threshold of membership. These communities usually have their own sets of rules and conventions that employ the use of colloquialisms specific to those communities. Additionally, they usually have a primary medium of communication between members. From this definition, she categorizes discourse communities by the context from which they arise, listing social, political, recreational, professional, and academic communities. Essentially, each community consists of people who voluntarily join a hobby or lifestyle, which in turn involves them within a discourse community.

Johns' discussion of discourse communities is the central lens for this paper's own discussions. Since people who watch YouTube videos have similar goals (to consume content), feature a common method of communication (comments), and meet the other four defining characteristics of discourse communities according to Johns, the YouTube discourse community can be rhetorically analysed and studied. Under this context, it can be understood that consumers of specialized YouTube content offer their own more niche community. This lens allows for research to build upon existing foundational perspectives for rhetorical analysis under the context of members acting within a larger discourse community.

Lu, D. (2018). Reddit's diverse platform: Transforming the social media landscape through its communities. *Stylus*, 9(2), 30-38. Retrieved from [https://writingandrhetoric.cah.ucf.edu/stylus/files/9\\_2/Stylus\\_9\\_2\\_Lu.pdf](https://writingandrhetoric.cah.ucf.edu/stylus/files/9_2/Stylus_9_2_Lu.pdf)

In his research paper, Daniel Lu explores the uniqueness of the social atmosphere on Reddit and how it can influence one's association to writing. He discusses that Reddit offers an environment conducive to social interaction and friendliness, since as he points out, the community promotes discussion. Additionally, since users are free to use any alias they desire, they do not feel the need to "uphold a public image" like other social media services. By sampling various parts of the Reddit community, Lu constructs the idea that writing appears in many forms, of which Reddit was one that made writing more enjoyable for him. It should be noted that Lu's paper centers around his own personal experience using Reddit, but does use tangible examples to illustrate his arguments. Though as a Reddit user myself, I believe this experience to be credible and paints a view of Reddit that can be translated to other social media services.

Lu's work offers a framework in perspective that can be useful in explaining the social atmosphere of other social media services, like YouTube. YouTube's community is certainly different than the Reddit community, but the aspects about YouTube's can be better explained in the context of other social media services to serve as a more holistic perspective. Similar to John's discussion of discourse communities and the aspects that define them, Lu describes aspects of communities on Reddit, which offers details specific to social media. In a similar way to Reddit, users are free to choose any alias to anonymise themselves if they so desire. This allows YouTube community members to act in a way without fear of tainting a public image or reputation. However, Lu's paper is only reliable to a certain extent. While I can vouch for the community and the social atmosphere on Reddit, the article essentially only describes the community in the context

of rhetoric. Lu, although being a member of some of Reddit's discourse communities, does not have a background or a very extensive experience in rhetorical analysis, being a freshman college student at the time of writing his paper.

Mitra, A., & Watts, E. (2002). Theorizing cyberspace: The idea of voice applied to the internet discourse. *New Media & Society*, 4(4), 479-498. doi:10.1177/146144402321466778

Anandra Mitra and Eric Watts propose that one way of looking at discourse on the internet is through the metaphorical lens of voice. By constructing a series of past understandings of voice and how it relates to physical discourse, they proceed to translate this to the discourse space of the internet, or as they refer to this: cyberspace. Essentially, since the internet has no central location or point in time, it contains all content from anywhere and all time. The internet's "potential to flatten hierarchies of power" is important to the modern context of internet discourse. From this, the authors mention the responsibility of internet speakers to act in a way that is proper. Since credibility can not be judged the same way as in real life, internet speakers' credibility is judged based on "emotional and ethical terms." Although this article depicts an outdated atmosphere of the internet, the fundamental aspects of speaking online has not changed, and therefore is relevant for discussion here. However, Mitra and Watts essentially write later that people who speak irresponsibly on the internet are subject to public review and therefore "it would be senseless for them to voice themselves on the internet." From the modern example of internet trolling and censorship, this aspect of the internet's discourse atmosphere has shifted.

This article offers an excellent “middle-ground” perspective between Johns’ lens of discourse communities to Lu’s observations of various Reddit communities. From using Reddit, it can be said with relative safety that Reddit represents a social discourse that more closely matches the overall internet discourse from 2002, as accounted by Mitra and Watts. The internet has become much more cynical, unbalanced in power, and ruthless in the last 15 years. Since this article was published, issues such as internet censorship and trolling have appeared as a result of the internet’s ability to anonymise and flatten power hierarchies. However, the article does allude to this possibility in the future when citing CNN’s credibility, even if not explicitly stated. Since this article represents a slightly older form of the internet, it can also be said that it represents a more *rudimentary* perspective of internet discourse. Since many additional layers of metasocial commentary have been added to internet discourse in recent years, it has become more difficult to identify fundamental pillars that construct how people use the internet. Basically, this article acts as an excellent “starting point” when transitioning from traditional discourse to internet discourse.

Sabatini, F., & Sarracino, F. (2019). Online Social Networks and Trust. *Social Indicators Research*, 141(3), 229–260. doi:10.1007/s11205-018-1887-2

Sabatini and Sarracino investigate the correlation between participation in social networking sites and trust. By utilizing data of face-to-face interviews with about 50,000 individuals provided by the Italian National Institute of Statistics, they analyse and extrapolate relevant information. After studying the data, Sabatini and Sarracino determine that “the use of [social networking services] is negatively associated to

people's propensity to trust strangers, neighbours, and the police." As not explicitly stated here, this also implies that trust with already trusted individuals are not significantly affected by the use of social media. The article is incredibly related to the central data source, that there is not much discussion of the implications for this data. It should be additionally noted that this data results from surveys during 2010 and 2011 in Italy, which may lead to some inconsistencies. For one, my research is centered around a Western perspective of the internet and may not match the same target demographic as those discussed in the other articles. Additionally, this particular paper is published eight years after this data (which may seem to be a short amount of time) but given the context of the internet's fast-paced environment makes the data questionable in accuracy. Although the aforementioned shortcomings, the study still appears to hold tremendous value since it does not explicitly rely on a causal relationship between using social networking services and their display of trust. This study shows simply that there is a correlation between the two factors, which can be more universally applied to modern situations. Within my own research and synthesis, this study may prove to be a helpful parallel between "real life" discourse and internet discourse. Including sources from demographics and perspectives outside of Western culture may show the universality of internet discourse. However with this said, the study does admit the theoretical aspect of itself. Since the interpretation of the data given is "merely speculative," there are many pieces of information that may be crucial in affecting the initial assumptions. The limitations of this study are limitations only in that not many concrete studies have been made of the effects of social media usage. The study, more than anything, shows that

more research and analysis must be done of the social climate resulting from the internet to understand it in its full capacity.