Variation in the Use of Twitter Hashtags

Allison Shapp
Qualifying Paper in Sociolinguistics
New York University
Spring 2014

Table of Contents

Introduction.................................................................................................................. 2
  History of Twitter and Hashtags ............................................................................ 2
  Tag Hashtags......................................................................................................... 4
  Commentary Hashtags ............................................................................................ 6
  Syntactic Variation in Hashtag Use......................................................................... 10

Background .............................................................................................................. 11
  Twitter and Hashtags ............................................................................................. 11
  Gender and Language ............................................................................................. 13

Methods: The Corpora .............................................................................................. 16
  Collection ................................................................................................................. 16
  Analysis .................................................................................................................... 19
    Independent Variables ........................................................................................... 20
    Dependent Variable ............................................................................................... 23

Quantitative Results ................................................................................................. 25

Tellability and Tweetability: Evaluation .................................................................... 33
  The Basic Narrative Structure ................................................................................. 33
  Tweets as Mini-Narratives? ..................................................................................... 35
    #evaluation ............................................................................................................. 35
    Internal vs. External Evaluation ............................................................................ 36

Discussion ................................................................................................................. 39

Summary ..................................................................................................................... 41

References ............................................................................................................... 42

Acknowledgements .................................................................................................. 44
**Introduction**

As a growing amount of everyday communication happens online, sociolinguists can identify and investigate new sites of variation in phenomena that don’t occur in spoken language but do occur in Computer-Mediated Communication (CMC). First we need to see what variants are possible, and then explore the semantic, stylistic, and other differences that one might convey by choosing one variant over another, by applying and adapting both existing quantitative and qualitative methods to new types of variables. This paper aims to start doing that with the linguistic tool of *hashtags*, developed on the website Twitter, and now spreading in use across the web. Hashtags are currently popular in the public eye, being voted Word of the Year by the American Dialect Society in 2013, and headlining several NY Times stories in the past few years.

Twitter is a medium used by millions of people daily and growing. In its Quarter 1 report of 2014, Twitter announced that it had 255 million monthly users, which was up 5.8% from the previous quarter (Shih 2014). Hashtags in particular, a tool originally innovated for the purpose of information organization and management, have gained many expressive functions. Moreover, they are a site of patterned variation, even as they are growing in use and changing in function. Existing literature focuses almost entirely on the original organizational tagging function of hashtags. That is, it focuses on the use of hashtags to label tweets with topics, in order to organize them and connect them to other tweets with the same label. In this paper, I call the canonically described genre of hashtags “Tags”, and explain and give evidence of non-tagging hashtags, which I refer to as “Commentary” hashtags. I give examples of several other ways hashtags can be used besides their original topic organization function, which I observed in a corpus of tweets I collected from Twitter. I do a quantitative variationist analysis on the frequency and context of use of both Tag and Commentary functioning hashtags, in which I find a strong effect of gender on the frequency of use of each type. Then I zero in on one specific function of the hashtag, evaluation, which I analyze from a discourse narrative perspective. Through this lens, I start to look for an answer as to why a Twitter user puts Commentary types of messages in hashtag form at all, instead of in plain text.

**History of Twitter and Hashtags**

Twitter is an online micro-blogging and social networking service. Users of the website can send and read short messages, limited to 140 characters, called “tweets.” Since the site’s launch in March 2006, it has grown to include over 500 million users (Semiocast 2012), is available in more than 20 languages, and is being used increasingly in many less widely spoken languages as well, as cataloged on www.indigenoustweets.com. Users can choose whether their account produces
private tweets that can only be seen by approved “followers,” or whether to make all their posts public, which can be seen not just by their followers but by anyone who visits the site, regardless of whether the viewer has a Twitter account or not. According to a survey by the social media monitoring software company Beevolve, almost 90% percent of Twitter accounts at the time the survey was taken were set to public (Beevolve, 2012). In April 2010, the United States Library of Congress acquired the corpus of all existing public tweets, and continues to archive all new public posts to the Twitter site (Raymond 2010).

An important tool available to Twitter users is the “#” symbol. Within the US and Canada, this symbol has historically been referred to as “pound sign,” or “number sign,” but it is largely referred to as the “hash symbol” in the rest of the English-speaking world. In addition, the symbol is and has been referred to as “hash” in many computer programming communities, including those within the US (Houston 2013). The most common usage of # across a variety of programming languages is as a mark for comments. It is a marker for meta-data, which is a set of data that describes and gives information about other data, basically “data about data.” Comments in code are written in the same document, but are not part of the code that will be run by the computer. Usually a # will eliminate the entire line of text that it precedes from being processed as code. Comments in programming usually explain what a section of code does, so that someone looking at it later can follow it better. Thus the symbol # has a history of marking meta-data, as it does also in hashtags, as I will describe below.

On Twitter, in a tweet, anything typed following the symbol #, until the next white space, becomes a clickable “hashtag” (the etymology being “hash” as in the symbol # + “tag” as in the act of topic-tagging information in order to organize it). Clicking on a hashtag brings you to a page that displays all public tweets that include that same string of characters. In this way, one can click on a hashtag and immediately read other tweets about the same topic. This allows one to follow and participate in discussions, which had been difficult to do before hashtags due to the organization of the website. Twitter users “follow” as many people as they want, and those people’s tweets will show up in the users’ “feeds” on their home Twitter pages. To find tweets not written by people you follow, one would have to search specifically for them. Hashtags provide a way to automatically find other tweets related to ones that you’re interested in.

The use of the hashtag on Twitter was first proposed by a Twitter user named Chris Messina (@chrismessina) in August 2007. He posted a tweet that said “how do you feel about using # (pound) for groups. As in #barcamp [msg]?” (Messina 2007) and followed up with a blog post explaining his proposal (Messina 2007b).
Hashtags were proposed amid a discussion to create user groups on Twitter to organize content, but Messina’s proposal focused on “improving contextualization, content filtering and exploratory serendipity within Twitter by creating a system of ‘channel tags.’” He was interested in “simply having a better eavesdropping experience on Twitter.” Messina proposed that hashtags be used as a means of organizing topics of discussion. The # symbol was proposed because Messina’s idea was modeled on existing conventions in a CMC medium called Internet Relay Chat (IRC), in regards to topic channels and the IRC community’s topic organizing philosophy and user-based implementation. IRC was already using the # symbol for public topic channels available across an entire IRC network, as opposed to those local to a server which were prefixed as '&.'

Hashtags on Twitter first gained traction in October 2007 during the San Diego bushfires. The event offered a clear use case for organizing discussion on a topic, and Messina was active in encouraging people to use hashtags to talk about the event. While hashtags were started and popularized by user initiative, they have now been incorporated into the infrastructure of the Twitter website. With their widespread use has come creativity and variation in how they are deployed. The hashtag was originally proposed for the purpose of tagging and organizing tweets to make having and following discussions on topics easier. This is the primary function of hashtags that has been focused on in academic literature, but in this paper I will illustrate another class of functions for the hashtag, Commentary, which I will give examples of and discuss in depth.

Since becoming popular on Twitter, hashtags have also been integrated into other social media platforms, most notably Instagram and Facebook. At first, just as with Twitter, hashtags on Facebook were a user-initiated convention that did not actually connect a message to others with the same hashtag. But during the summer of 2013, Facebook took the same step Twitter had taken in July 2009, and incorporated the hashtag’s function into the platform of the website. Since then, users have been able to click on hashtags on Facebook and see other posts using the same hashtag. This change whereby hashtags are officially incorporated into the program might have ramifications as to how they are used, but that is beyond the scope of the present study.

**Tag Hashtags**

The canonical type of hashtag, and the use it was originally proposed to do, is topic tagging. Tagging can be done to name a concrete entity, such as a person, place, company, or event. These types of Tags are often intentionally utilized for publicity, and can even be prescribed and promoted by a company as part of active advertising. Table 1 gives some examples of these:
The examples in the above table illustrate that Tags can be directed to different levels of the public. There are Tags that are relevant to tweets from the general Twitter public from all over the world, all over the country, and the like. These levels continue all the way down to tweets that are meant to tag entities but are only relevant for the individual author and her community of followers, or maybe even only the author herself. For example, #Newark is relevant to a large audience, while #VillageBelAir, an apartment complex in Los Angeles, is directed to a much more local public. Similarly, in events, #html5devconf is a developers conference attended and followed by people from all over the world. On the other hand, while #summerterm12 is very general and is probably used by many different small communities to refer to different entities based on whichever school they are attending, it is still salient as an organizing term to each author and her respective followers in a very local context.

The Tagging function also includes hashtags that are meant to organize tweets about a topic, which can be specific or vague. Here are some examples:

(2) To connect with others’ tweets about the same topic
   a) #PublicPrivacy is backed by the good faith of the American People
   b) Miss delacour = obviously not cut out for mazes. #harrypotter
   c) Cool shot of a 2011 eclipse: http://t.co/WOKw4x2f #alignment

(3) To organize among one’s own tweets, or those of one’s local network
   a) Just drove behind a truck full of chickens... our window and car are now covered in chicken poop #roadtripwithgene
   b) I’m not trying to be selfish or anything but all I’m thinking about is that next year this is gonna be me(: #ClassOf2013

(4) Circulating Memes
   a) #YouKnowYouGhetto if you arguing with your neighbor because he have a lock on his wifi!? Lol.
b) That was literally #TheBestDayOfMyLife! :)

c) #ThingsThatAnnoyMe when mfs talk badd about cameron(wiz) ! Dhatz alwayz

The examples in (2) and (3) show the same global/local distinction as we saw in the hashtags of entities. Tags can be meant to organize a national or international discussion on a topic, such as Public Privacy or Harry Potter, or they can be salient only within one’s own corpus of authored tweets, or those of one’s local network such as talking about one’s “road trip with gene” over the course of a few days, or referencing the “Class of 2013,” which could be used by many different communities to refer to different entities, but is expected to be able to be contextualized by the user’s followers.

“Memes” are a concept introduced by evolutionary biologist Richard Dawkins in his book “The Selfish Gene.” (1976). The word is meant to evoke the word “gene,” but comes from the Greek “mimeisthai,” meaning, “to imitate”. Dawkins coined the word to describe the concept of a “unit of cultural transmission” (Dawkins, 1976), an idea that circulates and replicates in society in a manner similar to how genes replicate.

In the context of Twitter, memes are common ideas in the form of hashtags that circulate and that are participatory, in that people learn about the hashtag and then use it themselves to add their own contribution to a funny or thought-provoking idea. The examples listed in (4) are lighthearted and meant to be funny, but memes can also be more serious. Recently there was a meme circulating with the hashtag #1reasonwhy, where women in the gaming industry (and some male colleagues) shared their experiences with sexism and discrimination in the field, explaining “one reason why” there aren’t more women working in the gaming industry.

Meme hashtags are a clear category of hashtags, but they complicate the tag/commentary dichotomy, as it’s possible that a meme could start as one person’s Commentary hashtag, but once it circulates enough and gets enough people to participate in using it, it becomes more Tag-like. Memes might follow a trajectory from Commentary to Gag. There are also memes that are more Gag-like and ones that are more Commentary-like. I introduce memes in the “Tag” section, but I describe more about how I dealt with them in my quantitative analysis below.

**Commentary Hashtags**

The hashtag was originally proposed as a tagging system, as a way to organize tweets about a certain topic. This is still the hashtag’s main function, and the one most talked about, but Twitter users have gotten more creative and started using hashtags for functions other than tagging. What I
call here the "Commentary" function of hashtags refers to hashtags that are used to add additional meaning to the main semantic content of the tweet, and are not intended to practically connect the tweet to others that use the same hashtag. What exactly this means will become clearer through the examples presented in this section.

The most common type of Commentary hashtag consists of those that give the author's evaluation about what they just said. This often takes the syntactic form of "Text of tweet in a full sentence. #evaluation," as illustrated in (5):

(5) Evaluation/Emotion
   a) Can this award ceremony start already... #anxious
   b) @TimFinch for real though .. How do I get it off ... Lol #annoying (;
   c) Why is Nelly's new song so catchy?! #loveit
   d) im tierd of being invited to these house parties for little kids and hookah bars for some “teen function” talkin bout some 16+ #yeahright
   e) @JaneSmith: That 3 second lap dance you get at the movies when someone walks by...
      #holarious

The misspelling (pun?) in Example (5)e underscores that this is not an earnest attempt to connect this tweet to others about "holarious" things. It is only meant to add evaluative meaning locally to the semantic content of the tweet. The hashtags of this type often name an emotion, but can also just indicate how the author feels about the content expressed in the message. For example, in (5)a the author indicates that they are anxious about the topic of their tweet, and in example (5)d the author expresses disbelief by tagging the tweet "yeah right," perhaps indicating disbelief that the guests are all actually 16+, but it's hard to interpret without more context. This evaluation type of hashtag will be analyzed in more detail below, using a discourse narrative framework.

Another site of creativity in the use of Commentary hashtags is situations in which the hashtag seems to actually be part of the main semantic content of the tweet, often an elaborate phrase, but is formatted as a hashtag rather than in plain language. The main difference between these and the evaluative hashtags described above seems to be that the main body of the tweets with evaluative hashtags stands alone, while the ones in the examples below either don't fully make sense without the additional content in the hashtag, or else the hashtag actually stands on its own as an additional idea, separate from the one in the main body of the tweet.
Part of the main content of the tweet

a) Had a dream that @LFarberrrr87 and I were in an all out battle against a heard of dear. #tookakicktotheheadthenwokeup #epic

b) Senior exit project. Microeconomics project. Math final. #onlythingslefttodo

c) @WassermanT Bate out from boozing. #youalwayscallwheneveryonesinbedwithahangover

d) @kperry34 Merp I could have some ice cream scooped by you any day ;)
 #needtogetmyassupthere

While I am not addressing here the motivation an author might have to include some semantic content as a hashtag rather than in a normal sentence, it is clear that the motivation of constructing these hashtags is not to tag the tweet to connect it to other tweets with the same hashtag. Many of these hashtags are extremely specific or complicated, and it’s very unlikely that many, if any, other tweets exist with exactly the same hashtag.

I also found several examples of templates that users can employ to be creative with their Commentary while still being part of a stylistic trend. One such construction I found in my corpora takes the form #team____. In the blank can go any entity that one wants to show an allegiance with. The “team” part is a metaphor, as the entity does not have to be related to anything that actually involves teams or competition. Below is a table with some examples of this construction.

<table>
<thead>
<tr>
<th>Full Tweet</th>
<th>Hashtag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good morning y’all. #teamnosleep lol</td>
<td>#teamnosleep</td>
</tr>
<tr>
<td>my name is Donnie Kimball and I say “AHAAA. Bro u guys cheer. Gtfo! Go suck a dick pus boy” Yes I am #teamGod.</td>
<td>#teamGod</td>
</tr>
<tr>
<td>People need to learn how to drive. It’s just rain people!! My truck almost ate a prius for breakfast!!#teamchevy</td>
<td>#teamchevy</td>
</tr>
<tr>
<td>when I tell your girl I’m #TeamTatted she fall in love</td>
<td>#TeamTatted</td>
</tr>
<tr>
<td>i make anything look good. i don’t dress for u I dress up for me. #teamAlicia</td>
<td>#teamAlicia</td>
</tr>
</tbody>
</table>

Table 2: Examples of tweets using the “team” construction

The examples in Table 2 show this construction being used to express affiliation with a range of different types of entities. While these uses of this formula are fairly disparate, they are all based on the same semantic relationship to “team,” and can be bent to the author’s creativity. I believe the formula to have begun in regards to rooting for one lover or another in a love triangle, such as in the Twilight Books. People were either “Team Edward” or “Team Jacob.” From there, the formula
spread. “teammnosleep” is aligning the author with the idea of having not gotten any sleep the night before, while “teamGod” may indicate that, despite the vulgar offering of the author, he is still a religious man, or, alternatively, it could be a display of super-egotism. “teamchevy” references a commercial product, and “teamTatted” refers to the group of people who have tattoos. Finally, “teamAlicia” indicates a stance where Alicia is showing confidence in herself. I chose to code hashtags of this type as Commentary hashtags, because they are expressing the stance of the author towards a particular idea or identity, and they seem to be largely idiosyncratic, created in the moment and not taken from an existing repertoire of possible “teams” to affiliate with.

As mentioned above, meme hashtags seem to be somewhat of a hybrid category, sometimes blurring the line between Tag and Commentary. Below are some examples of memes that seem to add additional semantic content to the tweet, and show the author’s evaluation and stance towards the main body of the tweet; as such, they function as a type of commentary hashtag. Examples are presented in (7) through (9).

(7) The purported “problems” of a certain group
   a) HootSuite is not working. #socialmediaproblems
   b) Sitting in Starbucks soaked from the rain in Midtown highlighting auditions in backstage waiting to go to work at 1230. #actorproblems
   c) I wish I could actively work in multiple songs at the same time. deciding which track to open is the hardest part of my day... #djproblems

(8) Extremely standardized ways of expressing a common emotion
   a) I have the worst job when it comes to my allergies #FML [Fuck My Life]

(9) Memes that become a template people use and add their own ideas to (as discussed above)
   a) #GrandTheftAutoTaughtMe i can go to the barber shop with a fade and leave with braids.
   b) #StupidThingsPeopleDo not turning on their turn signal.
   c) #IHATEWHEN PPL SAY I’M STUCK UP WHEN I’M NOT FOR REAL
   d) #ilostmyvirginityto nobody

The hashtags in the examples above often add an evaluation to the content of the tweet, but since they have become set sentiments, they circulate and can be used to organize discussion around that sentiment on a larger scale. It may be that any Commentary tweet has the potential to become a Tag if its use becomes widespread enough.

At first I viewed “meme” hashtags like these as a category of their own, distinct from Tags and
Commentaries. However, further examination of them convinced me that they too can be divided into “Tag memes” and “Commentary memes.” So in this analysis, memes were divided into the two main categories under study, but meme hashtags as a separate class could be made the main focus of a separate study.

**Syntactic Variation in Hashtag Use**

Another realm for variation in the use of hashtags in tweets is their syntactic role and position within the linguistic content of the tweet. The main distinction is between hashtags that are included in the syntactic structure of the sentences in the tweets and those that are extrasentential, most often tacked on at the end.

(10) Syntactic inclusion
   a) #Adderol is the only pill ill ever take...
   b) #iPayNoMind to those who say false things about me because i know the truth. Fuck the bullshit.
   c) @hahacraig I’m really looking forward to see your special tonight on #ESPNU. DVR is set so my daughters can watch as well!
   d) Free breakfast at #ikea

(11) Syntactic exclusion
   a) #NationalVegetarianWeek Twitter read my mind.
   b) Even husband of @amy jones is here! #EMCWorld Come to Hands-onLabs! http://t.co/lwQRhus5
   c) I’m going to be really upset if this game is postponed. #Yankees

The examples in (10) and (11) show what hashtags included in the syntax of the tweet look like and what those outside of the syntax look like. The examples displayed also illustrate that, whether included or excluded, hashtags can occur at the beginning, in the middle, or at the end of a tweet. Location of the hashtag within the tweet is one more way in which the distribution of hashtags can vary.

This section has described some of the variation observed in the use of hashtags on Twitter, both in their function and in their position and syntactic status within the tweet. I will now give some background on existing literature on hashtags, and also on the study of gender and language use on the internet, to set the stage for my findings discussed below.
Background

Twitter and Hashtags

In linguistics, Twitter has become a source from which to mine data, and studies have been done which use tweets to follow the diffusion of lexical items (Squires 2012) and the use of symbols such as emoticons (Schnoebelen 2012). Work has also been done on how Twitter can be used to study dialectal variation (Russ 2012). There are also studies of Twitter, and hashtags in particular, from a sociological perspective, including how Twitter is used for the organizing of (inter)national discussions on politics, crises (Heverin & Zach 2010), and the creation of publics through the organizing of topical discussion (Bruns & Burgess 2011). Within computer science and information science, there are studies that look at the mechanisms of how hashtags spread. These studies often use that data to feed machine learning models (Huang, Thornton, & Efthimiadis 2010; Chang, 2010). Previous work on hashtags, however, almost exclusively focuses on tweets as topic/theme hashtags. To date, there do not seem to have been studies on other functions that hashtags have come to serve. A search of the Journal of Computer-Mediated Communication, an open source journal available online, for the word “hashtag,” reveals three published articles, all following the spread of information on a certain topic through hashtags, similar to the sociological studies involving politics mentioned above.

One of the foremost Twitter researchers in linguistics is Michelle Zappavigna. In her article “Ambient affiliation: A linguistic perspective on Twitter (2011),” she begins by pointing out that the concept of searchable talk is something new, something that was not possible in canonical face-to-face interaction, but has been introduced in computer-mediated communication through the ability to tag talk with meta-data. Twitter hashtags are a form of in-line meta-data, meaning that they are integrated into messages rather than extraneous to it (as opposed to, for example, the hidden meta-data that is attached to a digital photograph about when it was taken and with what type of camera.) She argues that inline, searchable meta-data in online communication platforms like Twitter are effecting a cultural shift that lends itself to community-building activity. As Zappavigna puts it, hashtags function as a linguistic marker that says “Search for me and affiliate with my value!” She develops this attributive function of the hashtag and refers to it as “ambient affiliation.” The purpose of these types of hashtags is to organize and connect group discussion around topics and to affiliate both the readers and writers of the tweets with an ambient community of other Twitter users interested in the same topic. Zappavigna’s concept of ambient affiliation aptly describes the distinction between my Tag and Commentary hashtags. Tag hashtags seek to affiliate
with an ambient discussion or community, while Commentary hashtags do not, and rather function on a level local to the tweet, accomplishing more personal interactions.

Zappavigna uses a corpus comprising all tweets that contained the string “Obama” in the 24 hours after it was announced that Barack Obama had won the presidency of the US in 2008. She describes hashtags as “inscrib[ing] a keyword in a tweet as metadata referencing the topic of the message as assigned by the user.” She gives examples containing the hashtag #obama, where Obama is the subject of the rest of the text in the main body of the tweet. Her theory focuses on topic-tagging hashtags, which she describes as setting up a target that the content of the tweet serves to evaluate. In her corpus of tweets, #obama is usually the subject of evaluation of the rest of the text in the tweet. I propose that this characterization can be flipped around to accommodate Commentary hashtags in a related typology. Zappavigna characterizes the main body of the tweet as often an evaluation of the topic expressed in the Tag hashtag (like #obama). Commentary hashtags reverse those roles, with the hashtag instead being the evaluation of the main content of the tweet.

Zappavigna’s work also deals with an important discussion that is currently happening with regards to CMC: How can an online “community” be defined and described? Because of the text-based nature of the medium, community in CMC is necessarily created by language and discourse, making it a topic of interest for sociolinguists. Rheingold (1993) put forth one of the first definitions of ‘virtual communities,’ which he described as “social aggregations that emerge from the Net when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace” (Rheingold, 1993: 5). Work since then has sought to build on this definition in more concrete ways. The affiliative function of hashtags can be seen as one form of public conversation that, by grouping tweets, can create ad hoc social groups or communities.

Zappavigna concludes, and this paper will agree, that criticisms of Twitter as facilitating inane status updates about users’ activities have missed the social point of using CMC. “Twitter offers a medium for expressing personal evaluation to a large body of listeners with which one can affiliate ambiently” (2011: 803).

In a 2012 study of Twitter, Ruth Page makes a three-way comparison of the uses of hashtags by corporations, celebrities, and "ordinary" Twitter users. Page looks at Twitter as a means of self-branding and micro-celebrity, and particularly the ways in which hashtags serve to accomplish this. As this is largely accomplished through hashtags that create ambient affiliation, to use Zappavigna’s term, and link communities together using Tag hashtags, she focuses almost exclusively on that
type of hashtag. She does mention "evaluative" hashtags in passing (which by her definition encompass my entire "Commentary" category), but says that "hashtags are primarily used to make the topic of a tweet visible, rather than to emphasize stance. Idiosyncratic or expressive uses of hashtags do occur, but these examples are by far in the minority for this dataset" (2012:187). And indeed in her graph on page 188 she shows that among "ordinary" people, the rate of non-tag hashtags is just below 20%, for celebrities it is around 8%, and for corporations it is around 1-2%. I disagree that these "idiosyncratic or expressive" uses are so rare, as Page asserts. In my corpus, as illustrated below, I’ve found Tag and Commentary hashtags to be much closer to equal in usage – 42% of the hashtags in my corpus were Commentary hashtags, much more than Page found in her corpus in 2012, even among her “ordinary” users that most closely resemble my subjects.

One group Page focuses on is hashtags used by "ordinary" users that are for the purpose of self-branding, and promoting a professional self image. Her focus on corporate and celebrity use of hashtags to broadcast their endorsements while seeming to interact with their followers is concluded to be more like broadcasting and leveraging their audiences. Page’s results imply that it’s misleading to assume that the conversational qualities of Twitter are similar to typical dyadic exchanges. She says, "Fine-grained accounts of computer-mediated discourse are therefore needed to explore the development of this and other virtual marketplaces as a counterpart to large-scale studies of Twitter as a whole" (2012:199). The current study aims to answer that call.

**Gender and Language**

Gender variation in language use has been a focus of study in sociolinguistics since at least the 1970s, with many pointing to Lakoff 1975, Language and Women's Place, as an influential early work. Since then, there has been much research and discussion on the tendencies to be found in the different ways men and women use language, or at least the ways that feminine and masculine identities are performed in society through language. Much of this work has drawn on statistical correlations between gender and certain word classes. Lakoff proposed a set of features that has been the starting point for much research in this vein. She categorizes features ideologically associated with women’s speech in American English such as lexical items associated with women’s work, ‘empty’ adjectives (divine, charming, cute), tag questions, hedges, intensifiers, and hypercorrect polite linguistic forms. She describes the tendency of “women’s language” to be indirect while men are more direct.

One main way that gender-based variation has been described is in the use of language that is “informational” vs. that which is “involved” (also referred to as “emotional”) (Argamon et al. 2003). The “informational” type of language use focuses on propositional content, parts of speech
like prepositions and attributive adjectives, while the “involved” type focuses on creating interactions between speakers and their audiences (Biber 1989, Tannen 1982). The “informational” style has been correlated with males and the “involved” style has been associated with females, in studies of CMC corpora such as blogs, using predictive computational models (Argamon et al. 2003, Schler et al. 2006). This distinction has similarly been described as “formal” or “explicit” vs. “contextual,” using word classes such as nouns, adjectives, prepositions, and articles for the “formal” style associated with males, and pronouns, verbs, adverbs, and interjections for the “contextual” style linked to females. (Mukherjee & Bing Liu 201, Nowson, Oberlander, & Gill 2005). Research shows that these distinctions are not completely gender-based. Genre of writing often plays a large role in determining information or involved style, and so genres also take on gendered connotations because of these shared correlations. These language characteristics also come from the construction in western society of perceptions of appropriate emotion itself being constructed along gender lines, where men are expected to be more rational and women more emotional (Catherine Lutz 1986, 1990), and this affects ideologies and societal perceptions of what women’s and men’s languages are and should be (Hall 1995).

Of particular note is Herring and Paolillo (2006), a study applying the “informational” vs “involved” word class features to a corpus of weblog (blog) data. They collected a corpus of blogs that they classified into two genres, knowledge management or “filter” blogs, and personal journal or “diary” blogs. Their results showed that the diary entries contained more ‘female’ stylistic features, and the filter entries more ‘male’ stylistic features, independent of author gender. They also found that the author gender skewed in the same direction: more women write diary blogs and more men write filter blogs. Crucially, they found that the genre correlation with canonical gender-based features was much stronger than with the gender of the author. Tag and Commentary hashtags map very closely to the distinction made between knowledge management and diary blogs that Herring and Paolillo identified. Tag hashtags filter and organize data, while Commentary hashtags are more interactive and self-expressive, as the diary-type blogs are. I find a correlation in the same direction, with female identified users using the Commentary (“involved,” “interactional”) functions of hashtags more, and male identified users using the Tag (“informational,” “filter”) type of hashtags more.

Many studies have been done since Lakoff’s that use and build on the idea of categories or classes of words that can be shown to correlate with women’s or men’s speech. Most recently, this avenue has been pursued by computer scientists (and linguists) building predictive models that can use these correlations to predict the gender of the author of a piece of writing with great accuracy.
The Gender Genie formerly on bookblog.net, mentioned in Herring and Paolillo (2006), and based on an algorithm created by Argamon, Koppel, Fine and Shimoni (2003), was one such model that circulated as a novelty on the internet for identifying the gender of the author of literary or non-literary texts. It took into account language features that have come to be correlated with male and female speech, as well as the genre of the writing sample.

Many predictive models have been built using mainly word class features: Argamon et al 2007 obtained 80.5% accuracy of determining gender of the author with their model based on 19,320 weblogs, by using parts of speech, semantic classes of words, and other stylistic factors. Rao et al. 2010 assembled a dataset of Twitter posts from 1,000 users, and built a predictive model that combined n-gram features with more traditional word and phrase classes. Their best model obtained 72.3% accuracy. The accuracy attainable with these types of models shows that there are correlations between gender and language use that can be detected at least at the level of word class or n-gram frequency. But, of course, it is more complicated than just women behaving one way, and men behaving another, as the interference of genre as a factor has shown.

One of the most recent sociolinguistic studies to create a predictive model of gendered language use on the web also complicates the gender dichotomy by factoring in social network. Bamman, Eisenstein, and Schnoebelen (2012) built a model that predicted the gender of a Twitter user with 88% accuracy by using the canonical word classes associated with each gender, and adding a few new ones that apply specifically for CMC corpora (one of which was the use or not of hashtags, but this differed by only .02% between genders. I show that even if there is not a large difference in the rate of use of hashtags between men and women, they are used differently by each gender.) Bamman et al then looked at the 12% of users that their model got wrong and looked at the social networks of those users to construct a more complicated profile of gender and language usage on Twitter. They ultimately found that homophily in one’s social network, defined in the paper as “how much a social network is made up of same-sex individuals” (2012: 3), correlated more with the use of gendered language than an individual’s actual gender.

The advent of CMC studies has allowed for the study of people using language in an anonymous, disembodied environment. Huffaker and Calvert (2005), a study on teens’ use of blogs to build identity, say:

While physical constraints such as the body, biological sex, race, or age can have a profound effect on self-definition and self-presentation (Collins & Kuczaj, 1991), many of these attributes become flexible in online environments. In a virtual world, one even gets to construct one’s body. The anonymity afforded to youth within virtual worlds allows adolescents more flexibility in exploring their identity

15
through their language, their role play, and the personae they assume (Calvert, 2002)."

Despite this flexibility it has been shown that even in such environments, people often perform socially ingrained gender habits with regards to language, even though the unique anonymity of the internet could allow them not to. In Herring 1994, research is done to compare the behaviors of men and women in Usenet discussion groups, with a focus on adhering to norms of politeness (Brown and Levinson 1987). Herring finds that, while both genders profess in a self-report survey on netiquette not to like it when people flame, boast, or use profanity, in fact men perform these behaviors much more often than women do. This type of uninhibited behavior is often attributed to the decontextualized and anonymous nature of the Internet, but Herring found that this behavior presents differently in men and women, necessitating a factor of gender in explaining the difference in behavior between different users, not just the effect of anonymity. Herring’s theory is that, while women strongly follow the norms of politeness, men have other value systems that supersede politeness and license violations of both positive and negative face, even in this anonymous environment.

My research, while not a large-scale computational corpus study, goes beyond previous work by not looking at word or word-class frequency, but by looking more in depth at context, and observing variable use of pragmatic functions by the different genders by looking at types of hashtags used. I do not look at the frequency of words used in tweets by males or females, but the frequency of individuals’ use of the functions of Tag or Commentary hashtags. By finding a correlation between this behavior and gender, my work shows, supporting past work on gender in CMC like Herring 1994, that many users choose to perform gender even on the internet, where anonymity is possible. I will now describe my methods of data collection and findings regarding the use of Tag and Commentary hashtags.

**Methods: The Corpora**

**Collection**

The data in this paper comes from two corpora, a pilot corpus and a larger corpus, both collected using Twitter’s streaming API, which provides a random sample taken from all the public tweets being posted to Twitter in real time. The API allows large numbers of tweets to be downloaded at once, and takes parameters such as geographic location, to further filter the tweets.
I collected only tweets produced by users within the United States, and excluded any non-English tweets after they were collected.

The pilot corpus was collected on May 21, 2012, and the larger corpus on March 8, 2013. Both were collected using a utility called cURL to access Twitter’s API. The data is output by Twitter in a .json format. I then used a Python script (Patrick Callier wrote the original script, and I revised it to suit my purposes) to parse the data and extract relevant information.

The pilot data was taken from the three hours between 10am and 1pm EDT, and was limited to tweets within the US. The total number of tweets downloaded was 6765, and of those tweets, 993 (14.7%) contained the symbol # at least once. I used a sample of 500 of these tweets from the pilot corpus to form the basis for a preliminary account of variation in hashtag distribution and uses.

In the pilot study, I found a correlation between gender and hashtag function. Thus, in constructing a larger corpus, I collected a sample of users whom I could identify for gender, so as to be able to use gender as a sociolinguistic factor in my analysis. The descriptions and analyses presented here are based on these corpora as well as on participant observation. I have been the owner of a Twitter account since May 2011, and am a moderate but not avid user of the site.

To obtain the main, larger corpus used for this study, I first collected a random sample of tweets from a ten-minute time period starting at 7:25 p.m. on Friday, March 8, 2013. I chose a time after the normal work-day in order to get a representative sample that did not exclude those who cannot use Twitter at work. From the random sample of tweets, I looked at only tweets that contained hashtags. I then took all of the users who listed a first name (and had used a hashtag) and determined if the name was most likely a male or female name. This determination was made using US census data and gender probabilities for each name, calculated by Bamman, Eisenstein, and Schnoebelen (2012). If the name had above a .98 probability of being either male or female, then that user was selected for inclusion in the study. Names that had below a .98 score and names that did not appear on the list of the roughly 13,000 names included in the census data were eliminated for inclusion in the study.

Admittedly, this method did exclude a number of people from possible inclusion in the study. It excluded people with gender-ambiguous names, people who choose not to list their real names on Twitter (one of many motivations for which is not wanting to be identified by gender), and also people with unusual names that aren’t on the census list. However, because of the limited demographic information available about speakers on the internet, I believe that this is a reasonable way to start an investigation of gender and hashtag use.
The first 50 users with identifiably male names and the first 50 users with identifiably female names were put on an alphabetical list. Then the most recent 100 tweets (whether they contained a hashtag or not) from each user on the list were taken and assembled into a corpus of 10,000 tweets. From this corpus, a partial count of the tweets from 50 users (25 users of each gender) yielded far more tweets than were needed for assessing the function of hashtags, so the rest of the selected potential subjects were not used in this study. Retweets (RTs) were eliminated because they had not been authored by the users included in the corpus. The new total of tweets in the corpus was then 3,656 tweets. The total number of tweets with hashtags was 1072, or 29.3% of the tweets in the corpus, and the total number of hashtags was 1,633 (multiple hashtags can be used in one tweet.) This represents 1206 distinct hashtags, some used multiple times in the corpus.

The average number of tweets with hashtags in any given corpus is not a very telling number, as the distribution of hashtags is very uneven. Also, it must be remembered that users in my corpus were chosen only if they were shown to be hashtag users in the first place (having used at least one hashtag in the original random sample of the later corpus), so this set of users is biased towards being hashtag users. To better explore the distribution of frequency of hashtag use in my corpus, I assigned each user a “hashtag density” score, which was the number of tweets with hashtags divided by 100 (total sample of tweets for each person). These scores are shown in Figure 1, for the original 100 subjects. They show that most Twitter users employ between 11 and 40 hashtags per 100 tweets:

![Figure 1](image)

**Figure 1**
Although I took the same number of tweets from each user, the most recent 100 tweets they had produced, this covered very different time periods for each user. Some produced 100 tweets in several days, while it took others weeks or months to produce the same amount of tweets. These time periods are shown in the following graph, plotted by how many days it took each user to produce 100 tweets. I used three-day intervals for the first fifteen days and then larger intervals thereafter. About 1/3 of the users are in the 1-6 day range, meaning they write an average of 17 tweets a day.

![Days to 100 Tweets](image)

**Figure 2**

**Analysis**

I read each of the tweets in the corpus, and coded each of the 1633 hashtags within the tweets for the following linguistic and social factors, which I present in a summary table and then explain in more detail below. Note that the token in this analysis is the individual hashtag, and not the tweet, as a tweet can contain more than one hashtag:
### Independent Variables

**Internal Factor Groups:**
These factors are characteristics of the tweet within which the hashtag token occurs, as well as how the hashtag itself interacts with other internal elements of the tweet.

**Multiple:** This factor group refers to the number of hashtags in the tweet in which the token hashtag occurs (range: 1-11). Since the token of analysis is the individual hashtag, not the tweet, this factor is a measure of what kind of tweet a particular hashtag is occurring in, one with many hashtags or one with a single or few hashtags.

**Syntax:** The Syntax factor group measured whether the hashtag was part of the syntax of a sentence, outside the syntax of a sentence, or undetermined because the tweet does not contain any sentential material. See (10) and (11) above for examples of hashtags that are part of the syntax of the tweet and those that are not. The third value for this factor group, undetermined, is applied when there is no linguistic material in the tweet that the hashtag could possibly be a part of. This happens when the entire contents of the tweet consists solely of one or more hashtags, @ terms, and/or URL’s.

**URL:** The URL factor group indicates whether the hashtag cooccurs in a tweet with a URL, which is a link to content elsewhere on the web (e.g. an address like “www.twitter.com”). This could be a link to a website, a photo, or anything else on the web.

---

<table>
<thead>
<tr>
<th>Internal Variables</th>
<th>External Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor Group</strong></td>
<td><strong>Values</strong></td>
</tr>
<tr>
<td>Type of Hashtag</td>
<td>Tag, Commentary</td>
</tr>
<tr>
<td>(Dependent Variable)</td>
<td></td>
</tr>
<tr>
<td>Multiple</td>
<td># (range 1-11)</td>
</tr>
<tr>
<td>Syntax</td>
<td>part of syntax,</td>
</tr>
<tr>
<td></td>
<td>outside syntax,</td>
</tr>
<tr>
<td></td>
<td>tweet has no syntax</td>
</tr>
<tr>
<td>URL</td>
<td>Does or does not</td>
</tr>
<tr>
<td></td>
<td>co-occur with a</td>
</tr>
<tr>
<td></td>
<td>URL</td>
</tr>
<tr>
<td>@ [AT]</td>
<td>Cooccurs with an</td>
</tr>
<tr>
<td></td>
<td>@ term or doesn't</td>
</tr>
<tr>
<td>Position</td>
<td>Beginning, Middle,</td>
</tr>
<tr>
<td></td>
<td>End, None</td>
</tr>
</tbody>
</table>

Table 3: Summary of variables included in the quantitative analysis
**AT:** The @ [AT] factor group indicates whether the hashtag cooccurs in a tweet with an @, a symbol that refers to another user on Twitter and links them to the tweet (e.g. @gradstudent310 occurs in the tweet, where "gradstudent310" is a username belonging to someone other than the author of the tweet). This means that the tweet contained a term that directed it at another Twitter user directly. This might indicate that the author is talking about this other person, or they could be directing the tweet at another person in conversation.

**Position:** The position factor group indicates whether the hashtag occurred at the beginning, middle, or at the end of the tweet. There were a few occurrences where the hashtag was the only material in the tweet; in those cases, it is marked as having no position.

**External Factor Groups:**
These factors are characteristics of the author of the tweet that contains the hashtag token.

**Gender:** Gender of author (Male or Female). Gender was determined as described above; it was based on the probability that a name was male or female based on historical US census data.

**Friends:** The Friends factor group measures how many followers each user in the corpus has on Twitter (range: 21-1805, 17602, 28337). This is a measure of how big the audience of a Twitter user is. The user in my corpus with the fewest followers (subscribers) had 21. The range went into the thousands, with two outliers, users who had a very large number of followers, one with more than 17,000 and one with more than 28,000. I indicate these as outside the general range, because they are not typical, and indicate the user is some type of local (or bigger) celebrity.

**Hashtag Density:** This factor measures the percentage of tweets written by the author that contained hashtags (range: 1% - 91%) The value assigned to a given author was arrived at by dividing the number of tweets with hashtags (NB—not the total number of hashtags) by the total number of tweets collected for that user (almost always 100, but sometimes less if 100 tweets did not exist for that randomly selected user.) It is a measure, assigned to each user, of how frequently they use hashtags in the tweets that they write.

**Span:** Span measures the number of days it took the user who wrote the hashtag to author the 100 tweets that were in the corpus (range: 0.2 -456.1 days). This is a measure of how avid a Twitter user the author of the hashtag token is.
**Account Age:** Account age is the number of days the author’s Twitter account has been in existence *(range: from 808 to 2377 days).* I used a website that looks up the account age of public Twitter accounts to collect this data. At the time I did this, one previously chosen subject had deactivated their account, and so I could not find this piece of information for that one user.

The internal and external factor groups have different types of expected relationships to hashtags. Speaker gender is an external factor group, as gender is inherently independent of the Tag vs. Commentary opposition. I hypothesized that women would use hashtags to provide Commentary a greater percentage of the time than men. In contrast, the factor group of the presence or absence of a URL in a tweet fits more directly into the Tag vs. Commentary opposition. The relationship between +/- URL and Tag/Commentary is not causative. Rather, I hypothesize it to be correlative. A Tag (as opposed to a Commentary) and a URL have overlapping functions; they both refer to outside content on the web. This could be likewise true of @terms, in that they reference entities outside of the tweet, but always Twitter-internal, and internal to the network of the user. Therefore, one might not expect that correlation to be as strong as with URLs. With regard to the “multiple” factor group (a group sensitive to how many hashtags there are in the tweet), the more hashtags there are in a tweet, the more likely they are to be Tags. This is because of the nature of Tags and Commentaries. Tags are shorter, and listing a series of topic labels is a common practice for making something findable by keywords. Syntax and positions are overlapping predictors. The nature of a Tag makes it easier to insert into the syntax of the sentence; this leads to the prediction that many more Tags will be used sentence-internally than will Commentaries. Similarly, there is a clear hypothesis regarding the position of a hashtag in a tweet: Commentaries ordinarily act as evaluations of the main body of the tweet. As such, they are more likely to come at the end, and follow the sentiment that they are evaluating.

The external factor groups are all characteristics of the users who author the tweets that include the hashtag tokens. They largely do not have built-in expectations with regards to their correlations with uses of types of hashtags. In the case of gender, I hypothesize from the results of previous work that female users would tend to use more Commentary hashtags and male users would tend to use more Tag hashtags. For number of followers, I hypothesize that the more followers a user has, the more power using a Tag hashtag would have, and therefore the more they would be used. Hashtag Density, Span, and Account Age are all measures of a Twitter user’s habits with regards to how often they use Twitter, how often they use hashtags, and how long they have been a Twitter user. Together, they give a profile of a particular user’s entrenchment in Twitter,
which could predict a greater likelihood of atypical and innovative use of hashtags, hence avid users favoring Commentaries.

**Dependent Variable**

**Tag/Commentary:** The dependent variable is whether the hashtag fulfills the function of Tag or Commentary.

I read each tweet and used the context of the tweet (and the context of the author’s surrounding tweets that did not contain hashtags) the best that I could to determine whether a particular hashtag was fulfilling a Tag or a Commentary function. I used a conservative approach, coding hashtags as Tags, the assumed default, unless I could motivate them having additional meaning. One test that I used to make this determination was to think about if the person could reasonably expect someone to search for that hashtag, indicating it could be meant to index the tweet in a data organization way, to connect to a larger discussion or community, in other words, whether it could be reasonably expected to have ambient affiliation (Zappavigna 2011) or not. However, when things were meant to tag a topic, even if it might not connect to a much larger body, i.e. was only relevant on a local level as discussed above in Table 1, it was still coded as a Tag. Some examples of hashtags coded as Tag include hashtags for people, events, companies, television shows, etc. and also functional hashtags that do something within the program, i.e. crosspost to Facebook or Spotify.

One common format that I encountered and categorized as Tag is the “context template”, that looks like this:

```
Out of context statement in prose #context #context
```

Some examples of the context template from my corpus:

12) a) Crazy in Love #beyonce #superbowl #halftime
    b) Never curse or threaten me because there is never a reason to fight People. #OneLove #SeniorOlympics #Fight
When hashtags are meant to be a topic tag, even if it might not connect to a larger body, they are still Tags. That's why context is also a Tag, it still picks out a particular topic that the tweet is about, unlike Commentary hashtags. For the context hashtags, in this case the hashtags are the subject of the evaluation of the main body of the tweet, like in Zappavigna's theory, whereas with #commentary, it is the opposite way around: the hashtag is the evaluation of the body of the tweet. It's the stance towards what the author has just said, in that it is the method by which the speaker is creating and signaling a relationship with the proposition they wrote and with the people they are interacting with. (Johnstone, 2009; Hunston & Thompson 2000).

Commentary hashtags don't pick out a topic; they are messages within themselves. To make these fine determinations to the best of my ability, I searched the internet for context if I didn't understand something being said in the tweet, and clicked on all the URLs included in the tweets to see what was being referred to. If the hashtag added extra meaning to the content of the tweet, or if it was, or was part of, the major semantic content of the tweet then it was coded as Commentary. The sub-genre of evaluative hashtags (which will be focused on below in a discourse framework) were also coded as Commentary.

Originally I coded Memes as a separate category, but I came to see that even within the category of Memes, there were ones that were more Tag-memes and ones that were more Commentary-memes. So I broke them up, and used just the Tag vs. Commentary values for my dependent variable in the analysis. Multinomial analysis (doing mixed effect logit models with multi-level dependent variables) is also unreliable, because it is not straightforward to assign effects to the various variables, and effectively many models can fit the data equally well, making the interpretation of the results controversial. So ultimately a binomial analysis of Tag vs. Commentary hashtags fits best both the data and the method of analysis.

Twitter does not include punctuation as part of hashtags, so if a user includes punctuation in their intended hashtag, not knowing about this restriction, then Twitter will only make the portion before the punctuation into a recognized, clickable hashtag. I encountered only a few instances in my corpus where the writer clearly intended a hashtag that went wrong because they included punctuation. In these cases, I coded them as the intended hashtag. For example, if a user writes “#That'llPreach,” Twitter only recognizes the first portion “#That,” as the hashtag, or if they wrote “#L’amourParfait,” Twitter only recognized “#L,” as the hashtag, even though that’s clearly not what was intended. I coded for what was intended when it was obvious.

Finally, rather than guessing, I excluded a small number of tokens (N=6) that I could not make an informed decision on whether they were Tags or Commentaries with the given context.
that I had.

A general linear mixed effects regression model was fit in R, using the lmer package, and the results are in the table in the next section. The model included random effects for user and random slopes for position, and cooccurrence with URL and @ signs. A step down analysis was performed, resulting in the best-fit model listed below. The final model did not include Syntax as a factor group. Syntax and Position were very highly correlated, and therefore could not be run in the same model. Position is a more detailed predictor, as most ‘middle’ positioned hashtags are likely inside the syntax, but outside the syntax hashtags could be either initial or final. Ultimately, the model with Position had a better log-likelihood than the model with Syntax (all else the same), indicating that the model with Position was a better fit to the data, and it is presented below.

The dependent variable in the model is the type of hashtag, Tag or Commentary, so the model shows how likely a hashtag is to be either a Tag or a Commentary based on several factors, including information about the author of the tweet, and about other linguistic elements that the hashtag co-occurs with in the tweet. This is a study in how this kind of data can be approached with variationist methods, and find what sorts of variables are relevant in CMC data since we can’t use most of the regular ones, both social and linguistic. Often age, race, and class of a speaker in CMC are not available, and neither are phonetic or prosodic data when studying written speech. Since these are often the focuses of variationist studies of spoken speech, it’s necessary to adapt the methods. The field of CMC is still in the earlier stages of identifying relevant factors for the myriad new media of CMC that exist and are continuing to be invented.

Quantitative Results

Following are the results of General Linear Mixed Effects Regression (glmer) model. The effect sizes show likelihood that the hashtag is a Commentary, so positive effect sizes mean the effect is towards more Commentary use, and a negative effect size means the effect is towards more Tag use.

\texttt{glmer(TagComm~Mult+gender+friends+URL+AT+Span+HashDense+Position+gender:friends+friends:AT+HashDense:Span+(1+URL+AT+Position|screen\_name), data=d, family=”binomial”)}

\footnote{glmer(TagComm~Mult+gender+friends+URL+AT+Span+HashDense+Position+gender:friends+friends:AT+HashDense:Span+(1+URL+AT+Position|screen\_name), data=d, family=”binomial”)}
Cooccurrence with multiple hashtags in the same tweet is significant in the model (p<.001, Effect Size = .222). The more hashtags that were used together in a tweet, the less likely hashtags were to be Commentaries rather than Tags. While the number of hashtags in a tweet ranged from one to eleven, in fact, the preponderance of the data—93.5% of it—was in the range from one to four. This matches the hypothesis mentioned earlier based on the purpose of tagging as stacking keywords to make a tweet searchable. Also, Commentary hashtags appear to be longer and hold more semantic content, and so not as many of them can fit in the confines of a tweet. Tags are shorter and lend themselves to listing many in a row. Also, Commentary hashtags are separate ideas, so many of them in a row is a somewhat disjointed list of ideas, while the practice of listing topic tags is common and coherent.

Table 4: Results of glmer model

<table>
<thead>
<tr>
<th>Effect Size</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.240</td>
</tr>
<tr>
<td>Occurring with Multiple Hashtags</td>
<td>-0.222</td>
</tr>
<tr>
<td>Gender (Male)</td>
<td>-2.187</td>
</tr>
<tr>
<td>Number of Followers</td>
<td>-0.002</td>
</tr>
<tr>
<td>Cooccurs with a URL</td>
<td>-0.862</td>
</tr>
<tr>
<td>Cooccurs with an @ term</td>
<td>0.092</td>
</tr>
<tr>
<td>Span</td>
<td>-0.0241</td>
</tr>
<tr>
<td>Hashtag Density</td>
<td>-2.075</td>
</tr>
<tr>
<td>Position – End of Tweet</td>
<td>2.980</td>
</tr>
<tr>
<td>Position – Middle of Tweet</td>
<td>1.172</td>
</tr>
<tr>
<td>Position – Not applicable</td>
<td>1.557</td>
</tr>
<tr>
<td>genderM:followers</td>
<td>0.002</td>
</tr>
<tr>
<td>followers:@term</td>
<td>-0.001</td>
</tr>
<tr>
<td>Span:HashagDensity</td>
<td>0.0306</td>
</tr>
</tbody>
</table>

Signif. codes: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

Table 5

<table>
<thead>
<tr>
<th>Number of Hashtags in Tweet</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5+</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Commentary Hashtags</td>
<td>347</td>
<td>189</td>
<td>88</td>
<td>44</td>
<td>16</td>
<td>684</td>
</tr>
<tr>
<td>Total No. of Hashtags</td>
<td>718</td>
<td>433</td>
<td>234</td>
<td>143</td>
<td>105</td>
<td>1633</td>
</tr>
<tr>
<td>Percent Commentary</td>
<td>48%</td>
<td>44%</td>
<td>38%</td>
<td>31%</td>
<td>15%</td>
<td>42%</td>
</tr>
</tbody>
</table>

Table 5
Gender is significant in the model \( p<.001, \) Effect Size=-2.19). Males used more hashtags for Tag functionality and Females used more hashtags for Commentary purposes, as had been hypothesized. Here are the crosstabs and a graph of this effect:

![Commentary hashtags occurring in tweets with multiple hashtags](image)

**Figure 3**

<table>
<thead>
<tr>
<th></th>
<th>Tag</th>
<th>Commentary</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>347 (41%)</td>
<td>507 (59%)</td>
<td>854 (100%)</td>
</tr>
<tr>
<td>Male</td>
<td>602 (77%)</td>
<td>177 (23%)</td>
<td>779 (100%)</td>
</tr>
</tbody>
</table>

**Table 6**
This finding is similar to the result found in Herring & Paolillo (2006) for weblogs. Commentary hashtags are most similar to diary blogs, in that they express personal opinions and evaluations, while hashtags used for Tagging are more similar to the objective or generalized knowledge expressed in filter and knowledge-log blogs. This indicates that the subjects in my study, chosen because they list a gender specific name on their Twitter profile, are using hashtags in ways predicted by their gender. This paper does not present data to compare to the behavior of users who choose not to present a gender on Twitter, but the fact remains that a large portion of users choose to do so, and to exhibit gendered language behaviors, despite the available option to remain anonymous, or at least ungendered.

Cooccurrence of a hashtag with a URL was significant in the model (p<.0001, Effect Size=-0.862). URLs occurred less often with Commentary hashtags (only 15%) than with Tag hashtags (85% of the time). When there was no URL present, the hashtag was more likely to be a Commentary (56%) than a Tag (44%).

Including a URL necessarily entails a link to content outside of the tweet, elsewhere on the Internet. It indexes information outside the content of the tweet and outside of the discourse occurring on Twitter. As the Tag function involves affiliating with a wider community and managing knowledge, and is useful for a function on a level higher than the individual tweet, they may be used more with references to outside information (URLs) than Commentary hashtags, which serve a local function.
Span, the number of days over which a user wrote the 100 tweets included in the corpus, was significant in the model (p<.05, Effect Size=-0.024). The longer the span of time it took the user to write 100 tweets (the less avid a tweeter they are), the less likely they were on average to use Commentary hashtags, with the exception of the very most frequent tweeters, who did not use as many Commentary tweets as would be expected following the trend.

Table 7

<table>
<thead>
<tr>
<th></th>
<th>Commentary</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>no URL</td>
<td>599 (56%)</td>
<td>1065</td>
</tr>
<tr>
<td>URL</td>
<td>85 (15%)</td>
<td>568</td>
</tr>
</tbody>
</table>

Figure 5

Comparison of % of Commentary hashtag cooccurrence in tweets with or without URLs

Table 8

<table>
<thead>
<tr>
<th>Days</th>
<th>Commentary</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>185</td>
<td>43%</td>
</tr>
<tr>
<td>5-9</td>
<td>110</td>
<td>63%</td>
</tr>
<tr>
<td>10-14</td>
<td>175</td>
<td>64%</td>
</tr>
<tr>
<td>15-19</td>
<td>42</td>
<td>53%</td>
</tr>
<tr>
<td>20-50</td>
<td>107</td>
<td>41%</td>
</tr>
<tr>
<td>51-100</td>
<td>48</td>
<td>16%</td>
</tr>
<tr>
<td>&gt;100</td>
<td>17</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hashtag Density also came out significant in the model, with a relatively large effect size (p<.01, Effect Size=-2.075). This effect means that as hashtag density increases (meaning that as the percentage of tweets that a user writes that contain at least one hashtag increases), the user is likely to use more Tag hashtags over Commentary hashtags. Figure 7 shows this trend, as Hashtag Density increases, the percent of hashtags used that are Commentary (rather than Tags) trends downwards, although the trend is not very strong.
Finally, the Position the hashtag occupied in the tweet was significant. If it occurred at the end of the tweet, it was more likely to be a Commentary than if it occurred at other positions in the tweet (p<001, Effect Size =2.98). The vast majority of all hashtags, but Commentary hashtags in particular, come at the end of the tweet. The graph below shows the percentage of hashtags in each position that are Commentary, but note in the table that the data is skewed such that about 75% of all hashtags occur at the end of the tweet, and of those 1243 hashtags in the end position, they are roughly half and half Tag and Commentary. Tags far outnumbered Commentaries, percentage-wise, in the other positions.

\[
y = -0.0297x + 0.5404 \\
R^2 = 0.1741
\]

**Figure 7**

<table>
<thead>
<tr>
<th>Hashtag Density (Tweets per hundred that contain hashtags, by user)</th>
<th>Percent Commentary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-10% (N=57)</td>
<td>97%</td>
</tr>
<tr>
<td>11-20% (N=123)</td>
<td>77%</td>
</tr>
<tr>
<td>21-30% (N=303)</td>
<td>60%</td>
</tr>
<tr>
<td>31-40% (N=138)</td>
<td>49%</td>
</tr>
<tr>
<td>41-50% (N=291)</td>
<td>37%</td>
</tr>
<tr>
<td>51-60% (N=319)</td>
<td>31%</td>
</tr>
<tr>
<td>61-70% (N=29)</td>
<td>22%</td>
</tr>
<tr>
<td>71-80% (N=325)</td>
<td>17%</td>
</tr>
<tr>
<td>81-90% (N=0)</td>
<td>17%</td>
</tr>
<tr>
<td>91-100% (N=28)</td>
<td>17%</td>
</tr>
</tbody>
</table>

**Table 9**

<table>
<thead>
<tr>
<th></th>
<th>Beginning</th>
<th>Middle</th>
<th>End</th>
<th>Outside</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tag</strong></td>
<td>112</td>
<td>124</td>
<td>604</td>
<td>49%</td>
</tr>
<tr>
<td><strong>Commentary</strong></td>
<td>4</td>
<td>19</td>
<td>639</td>
<td>51%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>116</td>
<td>143</td>
<td>1243</td>
<td>131</td>
</tr>
</tbody>
</table>
A few other factors came out significant in the model but with very small effect sizes. Due to the number of factors I included in one model with this amount of data, those results are not necessarily reliable and won't be discussed in depth here.²

² Number of followers came out significant in the model (p<.01, Effect Size=-0.002), with a very small effect size, in the direction of people with more followers using more Tags. Interaction between number of followers and gender is significant (p<.05, Effect Size=0.002). The interaction between Span and Hashtag Density was also significant (p<.05, Effect Size = .031).
Tellability and Tweetability: Evaluation

We can use quantitative variationist methods to measure and explore patterns in how Twitter users use hashtags, but another question arises from looking at this data that needs to be explored from a different framework, and that is why users choose to use hashtags when and how they do. To begin explaining this, I liken a tweet to a narrative and a hashtag to the “evaluation” section of a story that makes it worth telling. In this section I go more in depth into narrative theory, and how it can be applied to tweets, and evaluation hashtags in particular.

Labov and Waletzky’s (1967) seminal work on the structure of narrative developed as a reaction to work that had been done on the narratives of practiced and professional story tellers. They were collecting sociolinguistic interviews from a variety of people in several different communities, and noticed the value of analyzing the narratives of what they called “unsophisticated speakers” (1967: 12). In their data, they saw similarities in the overall organization of oral narratives told about personal experiences, and also differences in the ways in which different people employed stories to express themselves. A basic structure emerged that framed the construction of narratives, but there was variation that reflected differences between individuals.

Labov and Waletzky stressed the need “to analyze the simplest and most fundamental narrative structures in direct connection with their originating functions.” (1967:12) Twitter is a medium in which people tell millions of very short stories every day. I argue that individual tweets, and groups of tweets published by a user in succession, can constitute units akin to narratives, exhibiting some key elements of the narrative structure described by Labov and Waletzky.

The main element that I focus on in this section is the evaluation component of a narrative. While included as a structural element of the frame of the narrative, the evaluation is a key functional element, giving the story meaning or validity as a “tellable” item. I found in a corpus of tweets that the hashtag is often used to express evaluation of the main body of the tweet, and contributes to the establishment of tellability in a similar way to that described by Labov and Waletzky. Below I outline the function of evaluation as defined in narrative theory and then provide examples of hashtags specifically fulfilling this function.

The Basic Narrative Structure

The basic structure of narrative laid out in Labov and Waletzky (1967) consists of five parts: Orientation, Complication, Evaluation, Resolution, and Coda. The orientation is optional, and is
made up of clauses at the opening of the narrative that orient the listener with respect to person, place, time, and behavioral situation. Orientation information can also be dispersed through the narrative, and is not always located at the beginning. The complication section is the main body of narrative material, usually comprising a series of events (or complicating actions). Labov and Waletzky define this section, and a minimal narrative, as being two narrative clauses ordered with respect to each other in terms of time.

After the main body of the narrative comes the evaluation section. Labov and Waletzky (1967) explain that a narrative consisting of just orientation, complicating action, and a result is not complete. Such a narrative is hard to understand because it lacks significance; it has no point. In a story of personal experience, the one telling the story uses evaluation to give the story meaning. An evaluation can do this in several different ways, for example, by tying the story back to a question that prompted it, by emphasizing the strange or unusual character of the situation, or by appealing to the element of mystery. By any of these methods, the evaluation is an explicit statement of the attitude of the narrator towards the story, and indicates the originating function of the story. It helps answer the question “so what?” and gives motivation to the telling of the story.

Often the evaluation is fused with the result in a single clause that both states the result and emphasizes its importance (35). The primary definition of evaluation is semantic, rather than structural. It is identified by the function it plays in the narrative above its structural form. The form of an evaluation can range in how embedded it is in the story (39). At the most embedded end of the spectrum, an evaluation might materialize as a symbolic action in the story or the evaluation of a third person. At the least embedded end, the narrator might make a direct statement to the listener about her feelings at the time.

Labov and Waletzky place the evaluation between the complicating action and its resolution (or coinciding with the resolution as mentioned above). An optional coda section comes after the resolution of the story. The coda is a mechanism for returning from the narrative world to the present moment. This can be accomplished in a number of ways, including through deixis (using “that, there, those” contrasting with “this, here, these” to show one is now removed from the story), or by following one character up to the present moment in a sort of epilogue.

This framework for narrative is the “normal form” that Labov and Waletzky found in their corpus of narratives collected during oral interviews. However, there is variation in the level of adherence to this normal form, in the optionality of certain elements, in the order of the information presented, and in the method of achieving the evaluation.
Tweets as Mini-Narratives?

I draw the analogy between narratives and tweets mainly from the shared functional element of evaluation. While historically, much sociolinguistic research on narrative has concerned the structure described above, more recently discourse analysts (e.g. Marra and Holmes, 2004; Ochs and Capps, 2001; Schiffrin, 1996) have turned towards a concern with narrative function (Vasquez 2007), which I believe tweets parallel, even if they don’t have all of the exact structural anatomy described by Labov and Waletzky. Many of the functions proposed for narratives, such as “to entertain, to express solidarity, to establish social relationships, to highlight and resolve tensions, to justify or explain one’s actions, to construct social and cultural identity” (Vasquez 2007) can all be accomplished in tweets and with the use of hashtags. In her research on workplace narratives by teachers, Vasquez highlights a maybe lesser recognized reflective function of narratives, a “‘making sense of reality’ function for the speaker.” (2007: 659). Tweets and evaluative hashtags in particular also can accomplish these functions, making them ripe for analysis as narratives in the functional rather than purely structural discursive framework.

Twitter is a way in which people express themselves and tell other people about events in their lives or thoughts they’ve had. It is a tool that people can choose to use, among many, to report on something that happened, just as narratives are. And also just like narratives, Twitter users employ evaluation frequently in order to give meaning to what they said, and to justify its being said. Some Twitter posts do convey a series of events in a temporal order, but some don’t. The ones that don’t still often report on events or thoughts in a way similar to longer narratives (although it is not to be assumed that Labov and Waletzky’s definition of narrative requires anything longer than a tweet, as two of the fourteen example narratives presented in their 1967 paper are indeed shorter than 140 characters).

Tweets reflect a new way that people “hold the floor,” in conversation with each other, that is worth exploring using some of the tools of narrative analysis. Labov and Waletzky use their description of narrative structure to show how “personal experience” narratives are different from other narratives. Twitter is a hotbed of expression of “personal experience,” and thus a worthwhile new venue in which to apply this apparatus.

#evaluation

One function a Commentary hashtag in a tweet can fill is the explicit evaluation of the content, or “complicating action,” of the tweet. While hashtags are found in all positions within a tweet, as discussed in the quantitative analysis above, one very common template I found was
tweets of the specific form: “Text of tweet as a number of clauses. #evaluation.” In these cases the hashtag is outside the syntax of any of the clauses in the tweet, tacked on as the last element, as illustrated in the following examples.

(13)

a) She out my life, can’t take it no more! #pissed
b) 30 minutes until lunch. #soexcited
c) Three conference calls today = very unproductive Monday. #sigh

Often the #evaluation is literally the name of an emotion, as in (13)a and (13)b above, but sometimes the evaluation is expressed through other means. For example, the evaluation can take the form of reporting a (possibly figurative) action like the #sigh in (13)c that expresses exasperation over the conference calls that are keeping the author from having a productive day. This evaluation of what she reported gives meaning to why she chose to post it. She is sharing this information because it is frustrating her and it is somewhat out of the ordinary. Just as in Labov and Waletzky’s evaluation, the evaluations in the hashtags here are the author’s voice, explaining what it is about the “story” that warranted its telling.

Explicitly saying the name of an emotion is a straightforward way of expressing evaluation, but it is a single strategy in the wide range of ways in which a hashtagged evaluation can support the tellability of a story.

(14) Do you remember as a kid taking jam from hotel’s free breakfast? Well I do; my little sister just did. #traditionscarryon

#traditionscarryon is a comment about the fact that the writer’s sister just did what she used to do as a child. The hashtag gives the author’s feeling towards the event that was just reported, and gives a moral to the story that explains why it was notable.

**Internal vs. External Evaluation**

In narrative theory, evaluation is divided into internal evaluation and external evaluation. (Labov and Waletzky 1967). Internal evaluation is evaluation of the content of the story, with the speaker evaluating the story from within the narrative world. External evaluation is an evaluation in regard to the interactional encounter. The speaker makes an evaluation not in the narrative
world, but in the present world about the action of telling the story. This distinction is present in the evaluations expressed through Twitter hashtags as well. In my corpus I found examples of both types, although instances of external evaluation were much more rare than internal evaluation.

The common template described above of a tweet followed by #emotion is a frequent version of internal evaluation. Some further examples of this are given in (15).

(15)

a) So excited that @Canoe93’s birthday gifts made it to Kent!!! #Yay #Excited
b) Got an A on my first history essay exam!!! :)) #SoHappy

The emotional hashtags in the above example, #yay and #excited, are commenting on excitement over the gifts actually arriving, not excitement about the action of telling people that the gifts arrived. The emotions are about the content of the utterance, not the act of uttering it, so they are internal evaluations.

An external evaluation in the context of a tweet is a hashtag that comments on the action of actually posting the information in the tweet. One hashtag that has this function is #TMI, which stands for “Too Much Information.” This hashtag occurred in my corpus twice, used by different users. (16) contains the two tweets that used this particular hashtag:

(16)

a) Someday I’ll get to walk around my house naked.. #TMI
b) I’ve been holding in my pee for awhile now. #Sorry #TMI

#TMI was used by these Twitter users to express that they think the information they just shared might have been beyond the bounds of what is considered permissible information; the tweets contain “Too Much Information.” (Still, they posted it anyway, with this disclaimer.) In these cases, the hashtag is not applied to the content of the post, but rather to the act of telling the information. The author of example (16)a is not expressing with this hashtag any evaluation of the idea of walking around her house naked, it’s the act of telling people (in a public tweet) that she hopes to walk around her house naked that she recognizes as “TMI.” The speaker here is offering an evaluation of her actions in the interactional encounter, rather than within the world of the story.

Any utterance has two levels of existence: one is the internal level of the information or story contained in the utterance, and the other is the external interaction created by the uttering of the utterance. Each of these levels can be evaluated by the speaker. We see through these examples
that tweets, just like verbal utterances, have these two levels of existence, and hashtags can be used to express evaluation of both of those levels.

All of Labov's stories are in response to a prompt (the majority respond to the “danger of death” prompt), because they were collected during one-on-one or group sociolinguistic interviews. Many times, the section in a narrative that Labov identifies as the evaluation is the clause or clauses that tie the story back to the prompt. Telling how the story relates to a prompt fulfills the evaluation function because it legitimizes the story as tellable in that situation. The story is licensed and relevant because it answered the question that was posed. The case with tweets is different, because there often isn’t a direct prompt that precedes the posting of a tweet. (Sometimes there is, though, when two people are tweeting directly to each other.) Most of the time, however, there is no direct stimulus that prompts the posting of a tweet.

However, the very fact of Twitter’s existence, in a way, is a prompt which invites people to create accounts and publish 140 character messages. There is a standing prompt which asks people to tell the public whatever they want to in this forum, and it is this prompt that each tweet is a response to. The existence of a Twitter user’s followers is also a perpetual invitation to say something, because there are people standing by who have signed on to “hear” it. In the current digital world, there are many prompts in a person’s daily life to publish their thoughts, including a number of social media sites, of which Twitter is one.

It is a common idea circulated today that modern social media users post frivolous things and share trivial information. Many of the short ideas shared in tweets like the ones presented here are criticized for being inconsequential. However, analyzed as mini-narratives, one can see that these (very) short stories are legitimized by evaluation through the same mechanism that canonical narratives are. They are also part of a “discursive construction of professional selves,” (Vasquez 2007).

Because of the differing nature of the prompt required to express oneself in the current digital world as compared to the 1967 situation in which Labov and Waletzky collected their data, perhaps differing levels or methods of legitimization are necessary to make something tellable. One interactional reason why legitimization of a story is required is that telling a story requires one to hold the floor longer than is normal in a conversation. In order to justify the extra floor time, the speaker needs to give a reason that the story is tellable, which is accomplished through evaluation. Again, the concept of “holding the floor” needs to be rethought in terms of computer-mediated interactional situations like Twitter. What does it mean to hold the floor, what social resources
does it consume, and what is required to justify the use of those resources? What does it take to make something not just tellable, but "tweetable?"

These questions remain to be fully answered with further study of media like Twitter as sites of linguistic and social interaction, but the analysis presented here shows that theories developed to analyze oral forms of expression can be fruitfully applied to new media, in order to illuminate the similarities and differences that exist in these new sites of interactional encounters, and that hashtags can be included among the “evaluative devices that contribute to meaning-making through narrative performance” (Trester 2013: 91).

Discussion

This paper has shown that there are at least two distinct pragmatic functions that hashtags on Twitter perform: Tagging and Commentary. One of their primary differentiations, focused on in this paper, is the intention of the author to participate in ambient affiliation (Zappavigna 2011). The original proposed purpose of hashtags was creating communities and discussions, building affiliation through language. But users have found new ways of using hashtags for interpersonal communication that also may build community, but among smaller communities of practice and on a more local level. Hashtags can be seen as being at the center of a discussion that is currently ongoing about what constitutes community in CMC, and about the new ways to build community that are created in these new media.

Reading through a whole corpus of tweets per user, rather than just one tweet per person, I came to notice some hashtag usage styles among the users in my corpus. For example, some people primarily use lists of Tagging hashtags, some people tend to use complicated and long hashtags of the “more semantic content” genre. Some people only use circulating Memes and Twitter functional conventions like #follow.

Many questions remain about the pragmatics of hashtags. In my corpus I encountered a small number of tweets that were just a hashtag: that was the entire content of the tweet. It is unclear with the amount of context I had in a corpus like this what exactly that is supposed to mean. There were also instances where a user hashtagged words separately that seemed to be one unit, for example

(17)  #Tiny #tornado strikes again http://t.co/M1dfUirY

“Tiny tornado” is how this particular woman refers to her toddler. Why not one hashtag
#tinytornado? What function does hashtagging #tiny and #tornado serve, when presumably the author is not wishing to connect her tweet (a picture of her son) to a wider conversation about tornadoes, or things that are tiny. I also encountered one tweet in my corpus where the hashtag seemed to be serving no function at all:

(18) @carleyann2312 @JT_bossin once again a trip for junk food #we be getting fat

It’s possible this was an attempt at a hashtag #webegettingfat that failed because spaces were added. It would be necessary in future research to have more contact with the Twitter authors themselves to gain more insight into all the reasons one might choose to use a hashtag rather than an alternative way of expressing the same thing. Further investigation needs to be more ethnographic and participant observation, as Eckert and McConnell-Ginet (1992) espouse about any research taking gender into account, that discussion of gender should be located within a particular community of practice. Even within Twitter, different levels of community of practice exist (small, local groups of friends, and arguably all the way up to the level of all Twitter users as a community), and it’s necessary to study local meanings attached to interactions and communications in these types of fluid CMC communities, and rethink the standard social-science definitions of community. Being able to more easily collect large corpora makes digital data very appealing, but sometimes it’s very difficult to interpret things like tweets out of context. One needs more information on the culture of the specific communities that are communicating with each other.

Commentary hashtags raise the question of why a user would choose to put that content in a hashtag versus just writing it in prose. What motivates someone to write a sentence and end it with #fact, instead of saying “This is a fact,” in prose? What’s the motivation for putting additional semantic content in a hashtag rather than as another sentence in the main body of the tweet?

My analysis of tweets in a narrative framework suggests at least one answer to these questions. Using a hashtag, you imply that something is more tellable, more tweetable specifically, and you show that you are part of the Twitter community and know the conventions. Whether or not it actually is very interesting, using a hashtag implies that the writer thinks what they are writing is “tweetable,” so they use the format and information structure that makes it specifically identifiable as a tweet.

Moreover, I have given evidence through finding a correlation between hashtag use and gender that social gender behaviors embedded enough that people perform them even when it isn’t unavoidable. They are doing it online as well, where there is no physical appearance or an
obligation to declare a gender. The subjects in my study chose to declare a gender, and chose to do canonically gendered things that matched those genders. I, as the researcher, don't even know that those genders are their real genders or that the names they provided were their real names, but the users who chose to list a clearly gendered name on their Twitter accounts, even when their physical gender presentation is not apparent, still seem to be performing behaviors socialized into those genders that they chose to present, namely acting more informational or acting more involved.

Summary

In this paper I have given a preliminary description of the uses of hashtags on Twitter. It is clear that they are currently being used for more than their original function of tagging topics and entities to organize discussion. Examples have been given of hashtags being used for tagging of concrete entities as well as topics, on both the global and local level. Among Commentary tweets, the evaluative role of hashtags towards the main body of tweets has been illustrated, as well as how hashtags are sometimes used to encode part of the main semantic content of the tweet. Twitter memes have been discussed as fulfilling both Tagging and Commentary functions on different levels, and at different times in the life of the particular hashtag. A quantitative variationist analysis was done that showed the performance of gender roles in the use of Tag and Commentary hashtags that is in line with previous research on gender variation in language use on CMC. It was also shown that there is systematicity in the use of CMC variables that don’t exist in spoken speech, but that are important to communication in new media such as Twitter. Evaluation hashtags in particular were looked at more deeply, using a qualitative discourse framework, to see how tweets are similar in structure to the canonical “narrative” and how hashtags can contribute to the “tellability” or “tweetability” of a tweet.

This paper has sought to identify the use of hashtags as a site of linguistic and pragmatic variation, and explore and describe some of the phenomena that occur, with the aim of supporting future research in computer-mediated communication, and specifically creative expression on Twitter.
References


Bamman, David, Jacob Eisenstein, & Tyler Schneebelen. 2012. Gender in Twitter: Styles, stances, and social networks. NWWA 41, Indiana University, October 27.


Messina, Chris (@chrismessina). “how do you feel about using # (pound) for groups. As in #barcamp [msg]?” August 23, 2007, 12:25 p.m. Tweet.


Rao, Delip, David Yarowsky, Abhishek Shreevats & Manaswi Gupta. 2010. Classifying latent user


Squires, Lauren. 2012. The lady pond beyond the lady pond: Variation in circulation, from TV to Twitter. Presentation at *NWAV 41*, Indiana University, October 27.


**Acknowledgements**

Thank you to Patrick Callier for help with data collection methods, and to Brice Russ, Lauren Squires, Matt Garley, Josh Iorio, and the members of the NYU CMC Seminar for help with the development of the research topic. Also, special thanks to Sean Martin for many helpful statistics consultations, Anna Marie Trester, and my adviser John Singler.