Laughing with Letters:

A Corpus Investigation of the Use of Written Laughter on Twitter

Isabel McKay
April, 2015
University of Michigan Linguistics Department
Senior Thesis

Advisor: *Deborah Keller-Cohen*
Second Reader: Robin Queen

1: Introduction:

> *"Their morning greeting to a friend in a distant city is usually "g.m.," and the farewell for the evening, "g.n." … The salutation may be accompanied by an inquiry by one as to the health of the other, which would be expressed thus: 'Hw r u ts mng?' And the answer would be: 'I'm pty wl; hw r u?' or 'I'm nt flg vy wl; fraid I've gt t mlaria.'"*

The passage above is taken from a New York Times article first published on November 30th, 1890 ("Friends They Never Meet: Acquainteneces Made by the Telegraph Key. Confidences Exchanged between Men who have Never Seen Each Other - Their Peculiar Conversational Abbreviations.," 1890). Its topic: off-the-clock communications between telegraph operators. The article describes how the men and women who worked the wires often became friends with operators in faraway cities; people who they never met. It provides examples of abbreviations like those above, but it also records several mythologized stories about long inter-operator feuds, describing both how operators were able to recognize one another by their typing styles and how they whined to one another over the wire.[1] A large portion of the piece concerns the ways in which these telegraph operators used language. To modern eyes the snippets of "telegraph speak" described in this article look almost exactly like the "txtspeak," language we use today on cell phones and online. In fact, the entire article bears an impressive resemblance to a Huffington Post exposé on the logistics of e-romance or a BuzzFeed piece on how kids these days can communicate using only emoji. Though numerous parallels can be drawn between the way operators seem to have used the telegraph, the world's first short-message service, and the way we use text and instant messaging today, the most compelling section of this article comes towards its conclusion.

[1] "Gol hang ts everlasting grind. I wish I ws rich"

> *"Operators laugh over a wire, or rather, they convey the fact that they are amused. They do this by telegraphing "ha, ha." Vary great amusement is indicated by sending "ha" slowly and repeating it several times, and a smile is expressed by sending "ha" once or perhaps twice. Transmitting it slowly and repeating it tells the perpetrator of the joke at the end of the wire that the listener is leaning back in his chair and laughing long and heartily."*

The passage above is fascinating not only because it describe systematic native speaker intuitions about the use of laughter over the telegraph, but the system described in the New York Times article is almost identical to the system described in similar articles written today about our present-day use of written laughter. An article posted to BuzzFeed in February of this year gives the following definitions between forms of *haha:*

> *haha*: "I'm acknowledging that you've said something you perceive to be funny, though I don't find it particularly funny myself,"
>
> *hahaha*: "That was funny! I legitimately laughed, or at least smiled, and I am slightly happier now than I was before you said that."
>
> *HAHAHA(etc.)*: "I am starting to panic that I may never stop laughing!!!!!!!!!!!!"

(Heaney, 2014a)

To modern readers it is likely not at all surprising that telegraph operators would have invented some means of expressing laughter over the wire, or even that their system might have been similar to our own. After all, what are our modern forms, *lol* (laughing out loud), *lmao* (laughing my ass off), *mwahaha*, **giggle**, and 😂 if not a version of •••• •– •••• •– (*haha*)? We write our laughter out every day and so it seems only natural that telegraph operators should do the same.

But why isn't it surprising? It's not as if *haha* can be found all through the novels, the essays, the newspaper articles, or even the personal letters that were written between the 1890s and the birth of the internet. It wasn't as though the first users of modern-day Short-Message Services (SMS) decided to type out their laughter in order to harken back to the days of the telegraph. Rather, each medium's convention of laughing with letters was adopted independently, because laughter is something we need in SMS communication today just as much as telegraph operators needed it in the 19th century. Even

today, with *lol* having found its way into most dictionaries, written laughter is more or less found only in short, personal communications. There is something about communication that makes us feel the need to use laughter and there is something about the way the telegraph and the text message force us to communicate in particular that has made us feel the need to type out that laughter.

This thesis represents a first attempt to understand the conventions and practices surrounding how and why we write our laughter with letters. In this project, an attempt will also be made to evaluate the extent to which use of written laughter draws on knowledge of physical laughter. The study of written laughter is complex, as it is both a phenomenon of short-message communication and an approximation of a face-to-face conversational tool. I will present evidence concerning the usage patterns of six forms of written laughter, *lol* (laughing out loud), *lmao* (laughing my ass off), *haha*, *hehe*, 😊, and 😂 on the social networking site Twitter.

The following sections are included in order to provide an overview of some of the previous studies of and observations on both face-to-face and written laughter which have informed this research. First, a literature review of previous research on both face-to-face and written laughter is provided. Afterwards, as there have been few studies examining written laughter, a section is included describing the history of each form under consideration and some native intuitions about each.

2: Literature Review

2.1: On the Study of Face-to-face Laughter:

Laughter is a human communicative universal (Edmondson, 1987). Though etiquette rules surrounding appropriate contexts for laughter may vary, the situations in which laughter arises are more similar across cultures than almost any other form of nonverbal communication (Glenn, 2003). Laughter

also seems to be one of our oldest communicative tools, as parallels of laughter can be found among several of the great ape lines (Provine, 2001).

Laughter is a deeply-engrained human behavior which can be studied through many different lenses. There is literature on laughter as physiological process, as a relic of human evolution, as a behavioral response to certain stimuli, as an indicator of emotional well-being, and as an identifier of humor. Though these bodies of literature are all important and interesting, this survey is limited to those studies which have examined the social information conveyed by laughter and its purpose in communication.

2.1.1: The Nature of Laughter

Early theories of laughter considered laughter to be little more than a side-effect of an internal psychological state.[2] Just as someone who sneezes can be inferred to have a cold, someone who laughs can be inferred to be experiencing the psychological state (whatever it may be) that accompanies laughter. Now experts tend see laughter differently. Modern-day conceptualizations treat laughter as a behavior actively intended to communicate how a laugher would like his or her words to be taken by co-participants (Glenn, 2003; Holt, 2013). These understandings draw on the fact that laughter heard out of context is only laughter. It may mean that someone is feeling nervous or that someone finds a joke

[2] Until about the last fifty years, laughter was treated in the literature more as a symptom of an individual's internal, emotional state than as a conscious communicative tool (Glenn, 2003; Provine, 2001). The assumption that laughter is a reflexive expression of a psychological state has been made most often in studies which describe laughter as a natural reaction to humor, but other theories have proposed it as a reaction to various other forms of stimuli. For example, Superiority/Hostility Theory argued that laughter was a reflection of an individual's emotions upon victory, a feeling famously described by Thomas Hobbes as "sudden glory" (Hobbes, 1640). Incongruity Theory, on the other hand, supported by Immanuel Kant and Schopenhauer, suggested that laughter reveals the shock a person feels when his or her expectations do not align with observed reality (Glenn, 2003). And Relief Theory, supported by Sigmund Freud and others, argues that laughter is a reaction to a release (or sometimes heightening) of some sort of psychological burden (Provine, 2001). Taken together these theories can fairly well approximate the types of situations during which laughter can occur. However, as shall be argued, there are some major flaws in their basic assumption: that laughter is symptomatic of a psychological state.

funny, but regardless, the presence of laughter does not necessarily indicate only one kind of emotional state. Most possible interpretations of laughter out of context have to do more with the surrounding interaction than with the feelings of the laugher; that a joke was told, that someone is threatening someone else, that someone just said something embarrassing. Though laughter, like other forms of nonverbal communication, is not regulated as consciously as language, it is still a socially-targeted action.

The most important piece of evidence that laughter is a marked act intended to communicate one's feelings, is that the presence or absence of others, whether or not those others are laughing, and the relationships between the laugher and the others in the room, are all important factors in the production or nonproduction of laughter (Glenn, 2003; Lee & Wagner, 2002; Osborne & Chapman, 1976; Provine, 2001). People are more likely to laugh in groups than alone, people are more likely to laugh in a group of friends than in a group of strangers, and people are more likely to laugh when others laugh with them than they are to laugh alone. Even, "canned" or "laugh-track" laughter can increase laughter production. Crucially, the variations described above seem to function independently from the perceived "funniness" of a situation. Most studies exploring the behavioral differences described above have asked subjects to view sitcoms or standup comedy routines. After the fact subjects were asked to report how funny they found the material they were asked to view. Generally, the "funniness" rating for individuals viewing the same content in various social settings was roughly the same, regardless of how much laughter was produced (Neuendorf & Fennell, 1988).

If laughter were a symptom of an internal psychological state, we might expect to see it inhibited in certain situations, as coughing can be inhibited in a quiet concert hall, but we would certainly not expect to see laughter inhibited in private. All of the environments in which laughter is most freely expressed are environments in which communicative channels to other participants are

opened wide. This suggests that laughter, like several other forms of nonverbal communication, is produced for the benefit of others.

There are, of course, situations in which individuals do produce laughter in private. Most theories place this sort of laughter production in the same category as talking to oneself or to the characters in a television show.

Rather than looking strictly to jokes and humor to understand the nature of laughter, most theories now suggest that an examination of a social interaction as a whole must take place. As feelings and desires are more or less unreachable for researchers, it is instead the context of laughter in conversation which must be examined. We must look at why an individual chose to place laughter when he or she did in conversation (Glenn & Holt, 2013).

2.1.2: The Meaning of Laughter

Conversation analysis (CA) is one of the most effective ways of looking at the effect various conversational actions have on the flow of an interaction. By examining actions and the precise locations where they occur in conversation, CA asks the question *why that now?* For this reason a CA approach to laughter is very enlightening. It asks, *why laughter here*?

An assumption of conversation analysis is that one of the major purposes of talk is to encourage intersubjective understanding between participants. This means that talk is used not only to communicate factual information but to communicate participants' opinions on those stances. This communication of opinion and feelings can even go as far as to express one's opinion on the ongoing exchange. Under this assumption it is not only *what* a person says that communicates information to other participants, but *how* and *when* that person says it. By replying to another participant's statement with a factually-relevant utterance an individual may display intersubjective understanding of the factual content of a previous utterance. However, by adding to that factually-relevant utterance, discourse markers, tonal inflections, or body movements, other sorts of intersubjective understanding can be

achieved as well. The speaker's level of comfort with the information being communicated may be understood by co-participants, the speaker's opinions on word choice in a previous utterance, the previous speaker's conversationally-expressed opinions, or the overall shape a conversation can be communicated through these small conversational actions. This suggests that by looking at utterances and interactional behaviors in terms of other surrounding, *adjacent*, utterances, an understanding of intersubjective meaning can be achieved (Peräkylä, 2007; Sacks, Schegloff, & Jefferson, 1974).

The groundwork for conversation analysis of laughter was laid by Jefferson, Sacks, and Schegloff in their 1977 paper, *Preliminary Notes on the Sequential Organization of Laughter*. Since this paper, several other similar projects have made important observations about laughter in conversation.

The first of these observations seems obvious, namely, laughter is indexical; it refers to something specific going on in the conversation (Glenn, 2003, 2010; Holt, 2011, 2013; Jefferson, 1984; Jefferson, Sacks, & Schegloff, 1977; Provine, 2001). Jefferson, Sacks, and Schegloff therefore suggest that laughter may be a *token of understanding*, which is a sort of conversational object which takes its meaning or force from its referent. The conversational object that laughter takes as its referent (generally known as a "laughable") may be either come before the laughter (as in a laugh response to a joke) or after the laughter (as in a funny story being introduced by laughter).

Second, most papers on laughter in conversation distinguish between conversational laughter, which comes in short bursts and does not drastically interrupt the flow of interaction, and "laughing together," which is raucous, extended, and involves most, if not all, of the participants (Glenn, 1989; Jefferson et al., 1977). "Laughing together" is often considered a separate activity in and of itself and is far less referential than conversational laughter.

Third, laughter is invited. We tend to think that it is rude to laugh at one's own jokes but, especially in small-group interactions, that is exactly what happens (Glenn, 1989, 1991, 2010; Glenn & Holt, 2013; Holt, 2013; Jefferson, 1979, 1984; Provine, 2001). In the majority of instances the person

speaking or creating a laughable is the first person to laugh. After the speaker's laughter has begun, others join in and "accept" the invitation. The terms "invited" and "offered" laughter are used, because it seems, at least outwardly, that the speaker's laughter gives other participants permission to laugh themselves.

Fourth, the genders of the individuals offering and accepting laughter seem to be important factors in determining whether or not an invitation to laugh will be accepted (Jefferson, 2004; Lampert & Ervin-Tripp, 2006; Provine, 2001; Rees & Monrouxe, 2010). Though several researchers have observed this difference, informed discussion on the topic has been restricted by the generally-accepted understanding that gender is performed, and thus any discussion of male and female behavior ought to be approached with care. Jefferson, however, observes that as long as these differences are discussed in terms of a "male identity projection" or a "female identity projection" (instead of male and female behavior) studying the laughter behavior of these groups can be legitimate. These studies more or less come to the conclusion that people projecting a male identity are more likely to laugh when not invited and are more likely not to laugh when invited, especially when the other participant is a female. People projecting a female identity, on the other hand, are more likely to laugh when invited, especially when the co-participant is male, and are less likely laugh when not invited (Glenn, 2003; Jefferson, 2004; Lampert & Ervin-Tripp, 2006; Provine, 2001).

The invitation/acceptance pattern of laughter is of particular note when attempting to determine the function of laughter in interaction. A laughter invitation asks a co-participant to interpret some laughable in a certain way, and a laughter acceptance indicates that the other participant has agreed on the interpretation (Glenn, 1989, 1991, 2003, 2010; Jefferson, 1979).

The leading understanding of laughter today is that is an indicator of a participant's desire to view the ongoing interaction through a playful or non-serious lens. This lens is known as a 'ludic frame,' (or more colloquially as a "play" frame) meaning that the social events taking place in and around

instances of laughter are to be interpreted as disconnected from reality; as a "pretend" version of a serious exchange (Glenn, 2003; Glenn & Knapp, 1987; Holt, 2013).

Play frames were first discussed by Bateson in the 1950s, though he did so using slightly different language than that we use today (Bateson, 1955). Bateson, in observing otters' play, realized that the actions involved were similar enough to those involved in fighting that the otters must have been indicating to one another through some sort of social signaling that the action taking place was play and not aggression.

Erving Goffman famously extended the ideas put forth by Bateson, developing the terminology of "framing" most commonly used today (Goffman, 1974). Goffman argued that by making certain signals animals like otters as well as humans are able to set up an interpretive frame of interaction. The same actions taken or words spoken in two different interactive frames may have wildly different interpretations. For example when a dog wags its tail, indicating a playful frame, and growls, other dogs will react by initiating a playful tussle, but if a dog were to give the same growl without signaling that a playful frame was in effect, other dogs might react aggressively or even violently. Humans, according to Goffman, signal frame in a similar way and for similar purposes, though Goffman proposes that our system of framing is more complex than a dog's or an otter's allowing for us to frame a wider variety of social contexts.

It is important to note that many behaviors intended to introduce a certain frame must be offered and accepted. This allows all participants to reassure themselves that everyone else is "keyed-into" the same frame (Glenn & Holt, 2013). For instance, if one dog growls playfully and another ignores her, a playful frame has been offered, but not accepted, and therefore a playful frame has not been ratified between the two participants in the interaction. Similarly if a one person laughs and the other does not (a situation often associated with hurt feelings) a playful frame has been offered and not accepted.

This definition of laughter, as a nonverbal cue implying willingness to enter into a playful frame, satisfactorily explains many of the puzzling observations researchers have made about laughter in interaction. First of all it explains why laughter is almost exclusively produced in social settings and is produced more often with a group of friends than a group of strangers. People do not need to signal play to themselves and people do not often wish to play with people they do not know well (Glenn, 2003; Provine, 2001). People may laugh when they are nervous or embarrassed in order to indicate that they are not taking an activity seriously and that therefore their failures need not be taken seriously either. Often a group of people will laugh when bullying a non-laughing victim; they will play with someone who does not want to be played with. People laugh when they find something funny so as to indicate both that they understand the playfulness of the laughable, and to welcome others to play with them.

Laughter in face-to-face conversation therefore provides illocutionary force, just like many other forms of nonverbal communication. By laughing the laughter indicates a desire for his or her words to be interpreted as play.

When we communicate over the internet, we need to communicate that play is underway just as we do face-to-face. If laughter is truly a marked, communicative act, it would not be surprising were we to find that the same work is being performed through some other means online. In searching for the conversational mechanism that indicates the offer and acceptance of play, the obvious first step is to look to written laughter. In the following section I will lay out some of the findings from the limited previous research which has focused on written laughter.

2.2: On Laughing with Letters:

The majority of the information which has previously been collected about written laughter was collected in studies focused on understanding other aspects of online communication. Studies which

have provided relevant information about written laughter tend to fall into two categories: studies of emoticons and studies attempting to identify user's "latent attributes" (gender, race, age etc.) based on language use.

2.2.1: Studies of Emoticons

The development of emoticons is, historically, very recent. Most accounts date the first emoticons back to 1982, when they were supposedly created by Scott Fahlman on a Carnegie Mellon message board (Krohn, 2004). It is unsurprising, therefore, that research on the topic has only just begun to mature.

An emoticon is traditionally defined as a series of ASCII characters (letters, numbers, punctuation, etc.) arranged in such a way that they seem to create either sideways or head-on images of facial expressions (Dresner & Herring, 2010). A few authors who have written on emoticons have also included certain emotive netspeak abbreviations (such as *lol* or *omg*) in this category as well (Krohn, 2004), but theirs is a nonstandard and largely metaphorical use of the word.

The standardization of Unicode has expanded the inventory of possible emoticons in most online environments to include single character images of facial expressions. Unicode also makes other single-character images available, such as trees, beer mugs, and small animals. Unicode images as a whole are known by the Japanese name for emoticons: emoji. Non-facial emoji are typically used either as decorative additions to a text message or else in making complete statements such as "let's grab drinks" (Bennett, 2014). The distinction between traditional emoticons (which generally, though not always, represent facial expressions) and emoji, especially non-facial emoji, is an important one to make, but as this is a subject which is neither well-studied nor particularly relevant to the topics under discussion, I will not delve into it here. In this paper I will use the two words interchangeably.

Etymologically, the word *emoticon* is a blend of the words "emotional" and "icon" (Dresner & Herring, 2010). The majority of work on emoticons has assumed, perhaps partly because of this name,

that emoticons are, essentially, depictions of a user's emotions; happy, sad, uncomfortable, etc. (Derks, Bos, & Grumbkow, 2008; Provine, Spencer, & Mandell, 2007; Walther & D'Addario, 2001). While it is true that often emoticons are used as stand-alone expressive acts, they are more often used to attach intent to a textual utterance, such as 'flirting,' 'joking,' or 'disapproving'. These intensions have been called emotions by some (Walther & D'Addario, 2001), but they are much more accurately described in terms of Speech Act Theory, as illocutionary force indicators (Dresner & Herring, 2010).

In their particularly insightful 2010 paper, Dresner and Herring discussed the illocutionary force of emoticons. They propose, in accordance with Social Information Processing theory of computer-mediated communication (CMC), that though emoticons may not map to facial expressions, they may accomplish similar communicative functions to those accomplished by smiles or winks (Walther, 2006). That is to say that while the same *information* may not be conveyed by a smiley-face that is conveyed by a smile, they convey the same *type* of information, namely they provide information about a speaker's intended meaning. Dresner and Herring also note that this type of information can be conveyed, to a certain extent, in more traditional textual styles either though textual statements of intent (ex: "just kidding") or through traditional punctuation (ex: "!" or "?"). They therefore claim that emoticons are a sort of internet shorthand which likely makes use of our pre-existing knowledge of nonverbal facial cues.

Though *lol*, *haha*, and other similar forms of online laughter do not qualify as graphical representations, and are not often treated as emoticons, they have been referenced as pragmatic particles which communicate illocutionary force (Curzan & Mejia, 2012; McWhorter, 2013). In other words, *lol* carries social information about speaker intent in much the same way as an emoticon might. One of the goals of this study was to establish the degree of similarity between different forms approximating the same face-to-face behavior, by examining abbreviated, onomatopoeic, and emoticon forms of laughter.

2.2.2: Studies of "Latent Attributes"

Though small studies and theoretical proposals such as those put forth by Dresner and Herring are evocative, the literature on emoticons is, as yet, not very extensive. The literature on other forms of laughter is even sparer.

Many studies of online communication focus on identifying what are referred to as the "latent attributes" of users (gender, race, age etc.) (Weller, Bruns, Burgess, Mahrt, & Puschmann, 2013). Such studies attempt to model the density and types of words and emoticons used by individuals of different groups in such a way that these groups can be identified by language use alone (Bamman, Eisenstein, & Schnoebelen, 2014; Lyddy, Farina, Hanney, Farrell, & O'Neil, 2014; Xy, Yi, & Xu, 2007). This information is particularly valuable to advertising companies that might wish to target a class of individuals in an anonymous setting. These quantitative studies, largely conducted on Twitter, have discovered that females use far more emoticons and "emotive" words, including various spelled-out laughter forms, than do males (Bamman et al., 2014). The work in this vein has been very successful in its practical goals, achieving greater accuracy in the identification of user gender than humans surveying the same data. However, in only a few of these studies have results been used to make theoretical claims about gender performance or gendered language use online (Bamman et al. (2014) being an exception).

Most word-frequency studies summarize trends in their results by placing word forms into categories, such as "emoticons," "numbers," or "hashtags." Though many of these studies mention forms of written laughter, none include analyses of "written laughter" as a separate category of analysis. Rather, these forms were typically separated into three categories; emoticon, abbreviation, and onomatopoeic (or slight variations on these). This means that *lol* and *lmao* are often lumped-in with forms like *bff* (best friends forever) or *omg* (oh my god); *haha* and *hehe* grouped with *AAAAA!* and *ouch*; 😂 and 😊 were treated as part of the same category as 😢 or 🏃. Though these forms are placed into

these categories for a reason, as generalizations *can* be drawn across these categories, this is the first work to have examined forms approximating laughter across these boundaries.

3: Laughing with Letters: Native Intuitions

The following two sections provide some information about the six laugh forms in question, *lol*, *lmao*, *haha*, *hehe*, 😂 and 😄 and about online laughter in general. Most of these laugh forms, which were chosen both because they are common and because they fit into several different categories of "emotive" netspeak expressions, are relatively new contributions to the English language. Where a project looking at older, more established, or more formal lexical items might turn to a dictionary or other "official" source for information on the history and meaning of a word, the sources available here tend to be informal and largely crowd-sourced. The major source used in this section is Urban Dictionary, an online, crowd-sourced dictionary of slang. While definitions on Urban Dictionary are crowd-sourced in that any internet user may provide definitions for words, several studies have used the site as a source for definitions of slang terms (Smith, 2011). Other sources of information include blog posts, humor articles, and other online media content. While the information taken from these sites is not academic in nature, it can provide native-speaker intuitions about how these forms are used. An individual may not think technically about an particular word, but so long as he or she is able to use it properly that individual has some sort of knowledge about the word's proper use.

The first section below presents form-dependent information, providing definitions for each of the laugh forms under observation taken from various sources. In the second section these form-dependent definitions are drawn together to outline some general intuitions about written laughter.

3.1: History and Practices for Six Written Laugh Forms:

3.1.1: *Lol*

The first documented usage of the initialism, *lol*, meaning "laughing out loud," was published in a 1989 FidoNet newsletter, though by that point the abbreviation had likely been in use in various online environments for some time[3] (Brandon, 2008; Edel, 1989; Hiscott, 2014; Pearson). As more people began using the internet, this term became both more active and more widespread. It has now even made the transition to spoken communication, where it is generally pronounced like the word "loll," rather than as an initialism (McWhorter, 2013). *Lol* has since been the basis of numerous lexical creations, including "lolzfest," "loller-skates," and the infamous "lolcats"[4] (Morgan, 2011).

The Oxford English Dictionary provides two definitions for *lol*. The first defines it as an interjection "used to draw attention to a joke or humorous statement, or to express amusement." The second defines it as a noun which is "an instance of the written interjection 'LOL'." ("LOL, *int.* and *n.2*," 2015).

The users who have posted about *lol* on Urban Dictionary over the years have made some other relevant observations about the meaning of *lol*. The quotations in 1 have been drawn from several users' definitions of the word.

1) **Urban Dictionary observations on / definitions of *lol*:**

   a. *"Now, [lol] is overused to the point where nobody laughs out loud when they say it. In fact, they probably don't even give a shit about what you just wrote. More accurately, the acronym "lol" should be redefined as 'Lack of laughter.'"* (no_one_2000, 2005)

   b. *"lol – originally meant "laughing out loud", but now is the most common expression in any text conversation, just used instead of HAHA or any giggle or something like*

[3] This abbreviation was also commonly used in handwritten letter-writing. In this medium, however, the abbreviation stood for "lots of love" or, according to some, "lots of luck." (Pearson)

[4] "Lolcats" being humorous photos of cats in costumes or silly poses subtitled using intentionally off-kilter spelling and grammar

that. Also used all the time when there's nothing else to say… LOL key should be added to a standard keyboard." (dude, 2005)

c. *"Lol is most commonly used as a silence breaker, a reply to a joke that is SUPPOSED to be funny but really isn't, or an answer to an uncomfortable or random statement that one couldn't think of a better response to."* (EXPLIZIT, 2005)

The use of *lol* is extremely commonplace in internet communication, so commonplace, in fact, that even by 2005 it had already lost a good deal of its power. John McWhorter has argued that *lol* has begun to be used as a more general "marker of accommodation," rather than as an actual indication of laughter (McWhorter, 2013). The perceived lack of connection between physical laughter or even humor and the use of *lol* is clearly an important part of its present-day meaning, and will be discussed at greater length below.

3.1.2: *Lmao*

Though the initialism *lmao*, standing for "laughing my ass off" is immediately identifiable by most users of the internet, it is neither as widespread nor as commonly used as *lol*. This initialism has therefore received less attention from dictionaries and in the media. *Lmao* and *lol* had similar origins in early online chat groups. A similar, though not identical abbreviation, *lmto* ("laughing my tush off") can be found described in the same 1989 article that represents the first written evidence for *lol* (Edel, 1989; Hiscott, 2014). Mike Vuolo dates the first recorded instance of *lmao* to a 1990 online Dungeons and Dragons game (Vuolo, 2013).

Unlike *lol*, *lmao* has not yet won a place in most formal dictionaries. But several Urban Dictionary definitions are given in 2, below.

2) **Urban Dictionary observations on / definitions of *lmao*:**

a. *"Used when the situation is considered funnier than a mere lol. I've also heard this said aloud in conversation pronounced: La mayo."* (TheDanishSugarfairy, 2008)

b. *"We use [lmao] when something's funnier and deserves more than a capitalized "lol" Sometimes people say it in person too, but that's only when they're trying to be funny"* (kirshteeen, 2009)

c. *"A chatroom acronym used exclusively by morons, meant to stand for "laughing my ass off". Generally typed in response to deeply unfunny remarks, and also used to feign nonchalance when taunting someone."* (Alan., 2006)

The word *lmao* seems to carry an implication that a comment was funnier than might have been implied had the user typed *lol* instead, but definitions like 2c still suggest a disconnect between typing laughter and actually laughing physically.

3.1.3: *Haha*

Unlike the two forms discussed above, both of which clearly trace their origins to online communication, *haha* is a word that has been around for thousands of years. Versions of *haha* can be found in various Indo-European languages dating back to Ancient Greece (Provine, 2001). The first record of it being used in English can be found in Ælfric's *Grammar* of Old English, written around 995 (Baron, 2011; Menzer, 2004).

All that the Merriam-Webster has to say about *ha-ha*, is that it is an interjection "used to express amusement or derision" ("ha-ha," 2015). This form is found in a large number of dictionaries, but always listed as two words (*ha ha*) or as a hyphenated construction (*ha-ha*). Some Urban Dictionary definitions can be found in 3.

3) **Urban Dictionary Observations on / Definitions of *haha*:**

a. *"Short way to let a person know over text that they are laughing/thought something you said was funny. It however doesn't really reflect how much they are actually laughing/how much they actually thought something you said was funny.*" (Entity1037, 2015)

b. *"Used to express laughter anywhere you can't say it, for example IM, instant messaging or via email. The minor difference between haha and hehe is that haha is often used when laughing at someone, while hehe is used while laughing with someone. It might be unconscious to most people, but it's true."* (Løkken, 2006)

c. *"A form of expressing laughter when oral expression is not available, like on the internet… Haha is almost never used to express actual laughter occurring. For those circumstances, rofl* [**r**olling **o**n the **f**loor **l**aughing]*, lmao, or a whole sentence about how the comment actually made one laugh is used."* (rikochet, 2008)

In addition to the intuitions evident in the quotations above some users in more recent definitions stressed the superiority of *haha* to *lol*. These claims tend to identify *lol* as juvenile. Others were careful to include specific and different definitions for various orthographic variations of *haha* as well.[5]

3.1.4: *Hehe*

Hehe is closely related etymologically to *haha*, as both are onomatopoeic. In fact, most formal definitions include *hehe* as a spelling variant of *haha ("ha-ha," 2015)*. The users who write definitions on Urban Dictionary, however, seem to see strong distinctions between the two forms.

4) **Urban Dictionary Observations on / Definitions of *hehe*:**

   a. *"muffled laughter, suggesting a sneaky aspect to that being laughed at, differs from lol in this way, which is a full on belly laugh.*" (monsieur_d, 2005)

   b. *"hehe, different from lol or haha. Hehe usually has some type of innuendo. It is a subtle way to flirt via texting or instant messaging." (guy12345, 2009)*

   c. *"A somewhat irritating giggle. You may find "hehe" pop up in various conversations in texts and ims. Many girls say this because it's a step cuter from the original "haha"."* (paperstars, 2009)

In general, *hehe* seems to carry both mischievous and diminutive or feminine implications. These implications separate *hehe* from more common laugh forms like *lol* or *haha.*

3.1.5: 😂 "Face with Tears of Joy": 1F602

😂, also known as "Face with Tears of Joy," or by its Unicode notation, 1F602, is the most commonly used emoji on Twitter ("Face with Tears of Joy," 2015; Rothenberg, 2015). This form was

[5] ricochet, the author of 3c, for example, gives the following additional definitions:
"Hah – The person thinks a comment is mildly funny.
Haha – The person does not think the comment is funny, but acknowledges your attempt at humor.
Hahaha – the person thinks the comment is funnier than average, or is just more enthusiastic than those who normally say *haha*
Hahahaha – The person is just being excessive. Usually used to strengthen a friendship because an inside joke was just mentioned."
Also see (Heaney, 2014a)

originally created to provide users with a means of expressing overwhelming joy. Specific information about the use of emoji is less freely available than is information about the spelled-out forms of laughter described above. This is partly because many websites like Urban Dictionary have not fully integrated Unicode into their system of operations (therefore not allowing users to created definitions for more unusual Unicode characters like emoji), and partly because the traditional dictionary model has not yet been extended to include non-word characters. Still, some blogs and magazines have published "How to Decode Emoji" articles attempting to describe the ways in which these emoticons are used and what they mean. Some examples of definitions for this form can be found in 5 below.

5) **The Meaning of 😂 (Various Sources):**

   a. *"LMAO* LMAO* LMAO* LMAO* LMAO* (*May not actually be laughing my ass off.)"* (Heaney, 2014b)

   b. *" **Translation:** 'I'm not as happy as I was when I sent the sobbing emoji, but I'm still happy.' … **Alternate uses:** When you didn't actually think something was funny and decided to use the minimal amount of effort to avoid hurting someone's feelings."(Toole, 2014)*

   c. *"LEGIT LAUGHING OUT LOUD"* (Sebastian, 2014)

Like many emoji, the use of 😂 in practice seems to differ considerably from its original conception. Rather than standing for "joy" this emoticon seems to be used as a form of written laughter. As with any Unicode character, the Unicode notation, 1F602, is expressed slightly differently on different operating systems and programs. On Twitter the 1F602 is expressed thus: 😂.

3.1.6: 😊 "Smiling face with Open Mouth and Smiling Eyes": 1F604

😊, also known as "Smiling Face with Open Mouth and Smiling Eyes," or by its Unicode notation, 1F604, is the first emoji to appear on most emoji keyboards. It is often considered to be a neutrally happy emoji. Though most meanings do not directly associate it with laughter, 😊 is often described as occurring in the same environments in which one might expect to find laughter.

6) **The meaning of 😄 (Various Sources):**

a. *"I'm not really into emoji, but I know you want me to be. This is the first one available."* (Heaney, 2014b)

b. *"The gigolo of emoji, this guy gets around our text chains, popping up to convey joy or sarcasm or to test your limits. Nothing comes across as too heavy when punctuated with* 😄." (Moss, 2014)

c. *"Hearty-u-really made me smile… This is as genuinely pleased as i can be on a particular day! So if am not dishing out many of those, well then that says a lot."*

This emoji has been included in this project because it falls somewhere between the traditional smiling face emoji and the 'Face with Tears of Joy' emoji discussed above. Like other emoji, 😄 is expressed differently on different systems. On Twitter it looks like: 😄.

3.2: General Observations on Written Laughter

The generalizations drawn in this section are based on observations made by native-speakers of an online "dialect." They may not prove completely accurate, but the observations made by native speakers can serve to point the way to substantive research questions.

First of all, several sources report that the different ways of writing laughter are not completely equivalent. Some are described as implying a certain sort of humor (*hehe* for example, implies flirty humor or innuendo). Others are described as being used preferentially in certain social environments (3b claims *haha* is used for 'laughing at' rather than 'laughing with'), or used primarily by certain people (2c and 4c both make claims of this sort). While different written laugh forms are similar in some respects, they also have some individuality.

Second, while most forms of written laughter are meant to imply that the user who typed the laugh form is physically laughing, many users report that they and others use them when this is not the case. This observation was made several times above for multiple forms (1a, 3a, 3c, and 5a). Appendix 1 also provides some examples of this observation being made in multimedia humor.

Third, while most forms of written laughter seem meant to imply that the user who typed the laugh form thought something in the previous tweet was funny, they are often used in response to remarks the user does not find humorous. This observation was made above for several forms (1a, 1c, 2c, 3a, 5b), and has also been made by several linguists (Curzan & Mejia, 2012; McWhorter, 2013).

Of these generalizations, the last two seem to make any association between spoken laughter and written laughter unlikely. If written laughter is neither universally associated with physical laughter, nor universally associated with humor, it seems as though the two cannot be accomplishing the same things.

The literature cited above, however, has made it fairly clear that laughter is not strictly associated with humor and that the purpose of laughter is communicative. Laughter is the means by which one expresses his or her desire to carry out a nonserious or playful interaction. When using the internet, expressing this desire out loud, through physical laughter, would be pointless. For all intents and purposes, an individual carrying on a conversation with others over the internet is sitting alone. Any physical laughter he or she produces would be akin to physically talking to co-participants far away and unable to hear. If one wants to communicate the desire for nonseriousness to co-participants who are not co-present the desire must be communicated through other means. It is thoroughly possible that the "other means" used on Twitter and other short message systems, from the telegraph to the text message, is written laughter.

4: Research Questions and Data

4.1: Research Questions

Is it appropriate to refer to *lol*, *lmao*, *hehe*, *haha*, 😂, and 😊 as written incarnations of laughter? These forms were originally meant to represent physical laughter, but might they instead be a *substitute* for physical laughter? This question has two prongs. The first is meaning-based. It asks whether or not

physical laughter and written laughter have the same meaning in a conversation. If face-to-face laughter indicates a desire to interpret a conversation as non-serious or playful, can written laughter do the same? The second prong is based in conversation analysis. This question asks whether the same rules govern the use of written laughter in conversation as govern the use of physical laughter in conversation. Do we see offer/acceptance patterns in written laughter as we do in face-to-face laughter? Are those patterns influenced by the gender of the conversational participants as well? Both prongs are essential to any argument that written laughter is truly a substitute for physical laughter in the written medium. This thesis, however, will focus on the second: on the placement of laughter in written conversation.

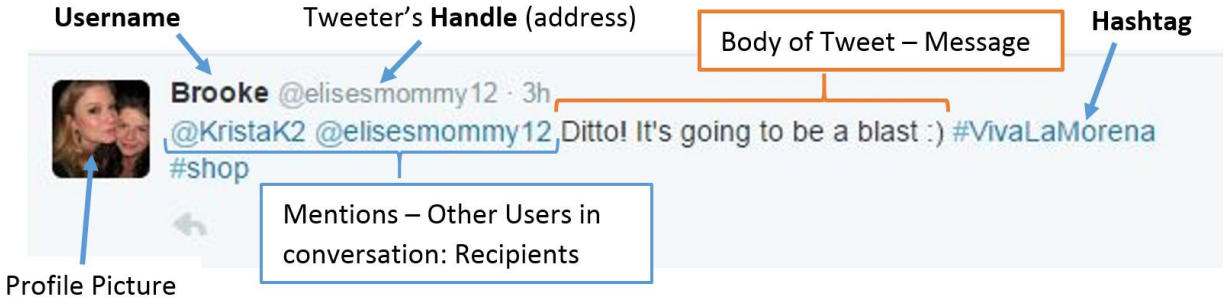
In order to establish whether or not we draw on our knowledge of laughter in interaction in order to use it in the written medium, it is first necessary to understand the ways in which laughter is used conversationally online. The analyses described below are all attempts to answer the question of how written laughter is used in conversation. As one of the goals of this project is to establish whether or not we draw on knowledge of spoken laughter to produce written laughter, these analyses are largely, though not universally, geared towards determining whether or not observations which have been made about the use of laughter in face-to-face conversation can also be made about the use of written laughter online.

4.2: Data

This project uses data drawn from the social networking site Twitter. Twitter is what is known as a "microblogging" site. This means that users create accounts which they use to post information about whatever they choose (personal life events, cat videos, news, recipes, etc.) completely publicly. Posts, limited in length to 140 characters, are known as "tweets," and may include various sorts of metadata tags. Tweets may include what are called "mentions," which use the syntax, @username, to

call another user's particular attention to a tweet. They may also use "hashtags," which use the syntax, #hashtag. Hashtags allow other users searching for tweets on a particular topic to find those tweets quickly, but they are also used rhetorically to provide meta-commentary on the underlying subject or meaning of a tweet (Yang, Sun, Zhang, & Mei, 2012).[6] Users may also include hyperlinks, images, videos, and various other forms of multimedia in tweets. All forms of text, however, count towards the 140 character length restriction. A diagram showing the layout and appearance of a tweet can be found in Figure 1.

Figure 1: Layout of a Tweet



Data were collected using Twitter's public streaming API (Application Programming Interface) ("Public API," 2013). An API is a "door" built into a website by its creators, which allows developers to use the site for their own purposes. Twitter's streaming API was constructed to allow developers to take and store limited amounts of data and access them in more flexible ways than are made available on the official Twitter website. Most major social networking sites provide APIs, but Twitter's is especially good for linguistics research because of the public nature of posting on the site. Other APIs, such as Facebook's, may allow researchers access to information about a user's friendship network, but their privacy settings are such that the text of posts or messages sent between users cannot be

[6] For example, in: @meagooon: "@Cynthiaaaxo LMAO I'm smiling thinking about it. Can't wait to love yew all day tomorrow in zeeee lib #studiousbitches", the hashtag #studiousbitches is likely not intended to help other users interested in "studious bitches" to find this tweet. Instead it provides a slightly humorous commentary on the tweet it is a part of.

retrieved (Weller et al., 2013). However, since posts made on Twitter are automatically public, Twitter's API can return information on language use not made accessible by other APIs.

There are several avenues, provided by Twitter and data-mining companies with access to Twitter's databanks, through which Twitter data can be accessed, but Twitter's free streaming API is the most effective for small research projects such as this (Weller et al., 2013). This API was used, for example in Bamman, Eisenstein and Schnoebelen's 2014 study of gendered language on Twitter (Bamman et al., 2014). This API has an effective "search" function, which allows a researcher to open a stream of all tweets containing a certain word, phrase, or set of words and phrases. From the time this stream is opened to the time it is closed these tweets are returned in real time, as they are posted to the site. In addition to returning the text of the tweets themselves, the API returns a wealth of metadata about the users involved in a tweet exchange. Most importantly for this project the API returns information about the tweeter's volunteered name, username, and location, a Twitter-provided timestamp, and information about whether or not the tweet is a retweet, whether or not the tweet was a reply, and a certain amount of user information any users mentioned in the body of the tweet ("The Streaming APIs: Documentation," 2014; Weller et al., 2013).

The free streaming API is intended for small-scale use only, and therefore suffers from many of the same limitations as most forms of free software. The limitations of the free streaming API are as follows. Firstly, all APIs are limited by the data provided to the company by users upon subscription. If a user wants to give a false name or a false location, Twitter allows them to do this. Therefore though the API can allow one access to who and where a tweeter *claims* to be, this information cannot be taken as fact. In many instances users will state that their location is, to name a few examples, "at Hogwarts" or "lost in the now." Users of Twitter are asked, but not required, to state their gender, and are not asked to state their racial background upon registration. Gender and race information is not available through the streaming API. Secondly, because one of the major ways that companies like Twitter turn a profit is

by selling data, they limit the amount of data that can be extracted using the streaming API. The free streaming API only allows a user to channel 1% of the total information flow on Twitter at any given time. Higher-level permissions can be purchased, but were not necessary for this project. If the 1% cap is, at any point, exceeded, Twitter simply returns a random sample of tweets matching the search criteria ("The Streaming APIs: Documentation," 2014; Weller et al., 2013). As a random sample is all that was needed for this project, this cap did not present an issue for data collection.

Initial data collection took place in November of 2014. Using the streaming API, and a filter which removed automatic retweets, 1000 examples of tweets containing *lol*, and 250 examples of tweets containing *lmao*, *haha*, *hehe*, 😂 and 😊, were collected in samples extracted five times over the course of one Saturday. Initially the dataset for *lol* was larger than the datasets for the other five forms under observation because *lol* was to be the sole topic of this paper, with the other forms merely providing context. Once the analysis began, however, the concept of this paper began to shift and it became necessary to conduct additional collections.

These collections were necessary for two reasons. Firstly, a few of the analyses in the sections which follow require large datasets in order to provide statistically powerful results. In order to provide comparisons between *lol* and some of the other laugh forms under observation the datasets for *haha* and 😂 were expanded to be an equal size. Due to time constraints and the amount of effort required to hand-tag datasets, the samples for *hehe*, *lmao*, and 😊 were left at 250 and were simply left out of some of the comparisons. The second reason for additional collections was that the initial collection had not provided a comparison group: a random sample of tweets with which to compare some of the traits of tweets containing written laughter were emerging. Additionally, some of the analyses required a comparison group of tweets specifically *not* containing laughter.

A random sample group of 600 tweets (a number large enough to provide powerful results for the analyses which needed the sample) was therefore collected in one block on a Monday in January of

2015. Also, an additional 750 examples of tweets containing *haha* were collected in one block on a Wednesday in January of 2015, and an additional 750 examples of tweets containing 😂 were collected in one block on a Wednesday in February of 2015. Though this collection of tweets on varying dates was not ideal, it does not challenge the validity of this study. The difference in collection dates was only a few months, likely not a long enough time for the use of forms as established as these to undergo major changes in usage. The goal of this project was broadly to understand the ways in which these laugh forms are used on Twitter, and these late collections are still able to serve that purpose.

The result of these collections was seven distinct datasets, an unfiltered sample of tweets, and one dataset of tweets containing each of the six forms in question. After examples for which the desired form was included in a manual retweet rather than in the new contribution portion of the message were filtered out, the resultant datasets were as shown in Table 1. Retweets both manual and automatic[7] were filtered out of the datasets because they could not be either attributed to the original posters or located within their original conversational context. As tweeter-identity and conversational contexts were both relevant variables in the analyses below these tweets could not be used. However, those tweets containing manual retweets for which the dataset laugh form rested in the part of the tweet which did not constitute a quotation were used. These will be discussed in greater detail below.[8]

Table 1: Dataset Sizes

	Haha	Hehe	Lol	Lmao	😂	😄	Unfiltered
Target Size	1000	250	1000	250	1000	250	600
Resultant Size	984	349	963	248	992	240	594

In addition to the text of each tweet, the user-provided name, user "handle" (address), and location along with the timestamp of the tweet were automatically added into the dataset. There was only one instance, in the *lol* dataset, in which the same tweeter contributed more than once to a

[7] Recall, however, that automatic retweets were filtered out on collection.
[8] See Section 5.4

dataset. The second instance was removed. This collection therefore provides data about broad usage rather than the repeated practices of a small set of individuals.

In some of the analyses below it became necessary to filter these datasets still further. Most commonly, the unfiltered dataset was reduced to exclude tweets which contained any form that might possibly represent written laughter, creating a "no laughter" dataset. Where the laugh forms were spelled-out with letters the identification of these forms was simple. The forms which are considered spelled-out laughter throughout this paper are: *lol*, *haha*, *lmao*, *lmfao*, *jk*, *hehe*, *hoho*, *ahah*, *huhu*, */chuckle*, **giggle** and orthographic variations of these. Identification of emoticon forms of laughter was more difficult. As there is no set list of emoticons which can and cannot be used to represent laughter, broad identification mechanism which likely over-identified forms was used. Any emoticon form with an upturned mouth was considered emoticon laughter. Throughout this paper the following forms, along with ASCII variations on these, will be referred-to as "emoticon laughter": 😊, 😍, 😂, 😄, 😎, 😉, 😆, ☺, 😃.

Often the laughter datasets were filtered down only to examples of what will be referenced here as *isolate laughter*. Tweets containing isolate laughter contain one and only one instance of the dataset laugh form and no other instances of either emoticon or spelled-out laughter. Instances of isolate laughter are desirable for two reasons. First, many of the comparisons which follow will make comparisons between the behavior of individual laugh forms. If more than one laugh form is present within a tweet these other laugh forms may interfere with data gathered for these comparisons. Second, many of the comparisons which follow use the position of a laugh particle within a tweet as a variable. If two laugh forms exist at different locations within a tweet this might similarly interfere with results. Tweets with isolate laughter constitute over 90% of the total sample.

In addition to the intra-utterance usage of online laughter words, a major concern of this project was the conversational context in which written laughter could be found. Among other things, this

context provides opportunities to seek out the offer/acceptance patterns seen in face-to-face laughter. That being the case, it was also necessary to collect any conversational context surrounding the use of a laughter word. Though Twitter's API has the ability to collect a certain amount of information about the tweet directly preceding a particular tweet in context, it cannot provide information about subsequent conversational turns, as tweets are collected in real time. Subsequent context must therefore be retrieved at a later date. Conversational context for each tweet was collected some time after major data collection, using the "view conversation" function on Twitter's website. As the layout and profile pictures associated with each user was also desired for reasons discussed more extensively below, conversational context for each tweet was collected as a screenshot. Though many of the tweets found in these conversational contexts did contain laugh forms, these context tweets were added to the original datasets only as they pertained to the original data collections: as metadata. Each tweet was therefore connected to a single screenshot of its conversational context, and each screenshot of a conversational context was connected to a single tweet.[9]

These seven datasets were tagged for several variables, and these variables were analyzed in terms of frequencies. As each method of tagging is specific to an individual analysis, these tagging methods are described analysis by analysis in the sections which follow.

5: Analyses and Results

The following sections represent the various analyses which were carried out using the datasets described above. Section 5.1 examines the overall frequency with which various laugh forms occur on Twitter. Section 5.2 examines the frequency with which various laugh forms co-occur with three of the forms under observation (*lol*, *haha*, and 😂). Section 5.3 examines the influence that the presence of

[9] This will be discussed in greater depth in the introduction of Section 5.6 below

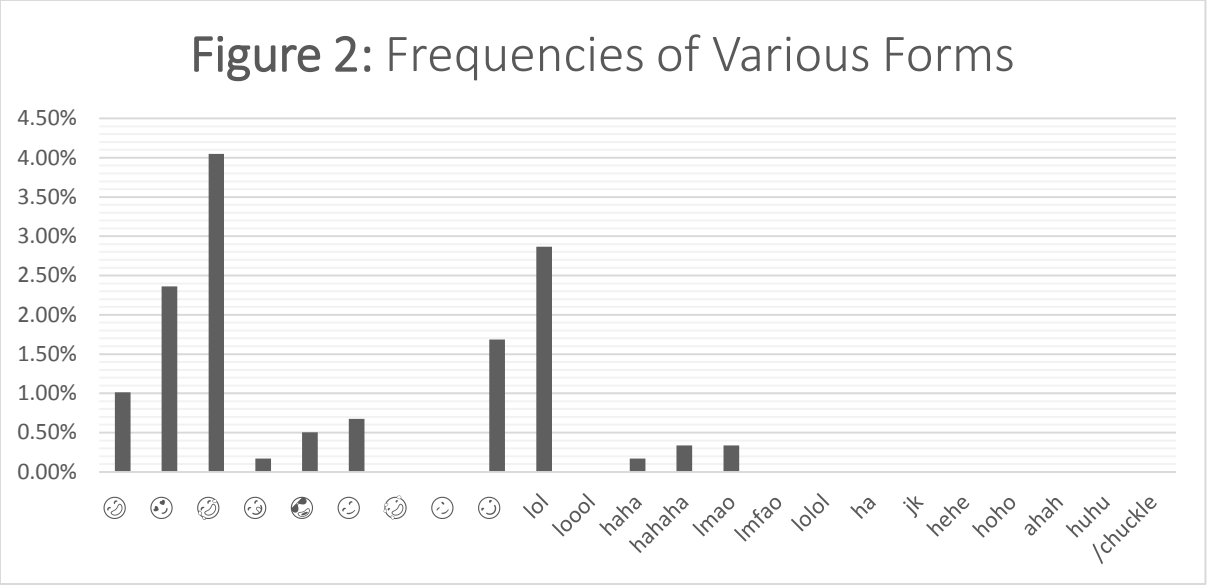
written laughter within a tweet has on that tweet's likelihood of being designed for a specific recipient. Section 5.4 models the placement (initial, medial, final, alone) of laugh particles within a tweet. Section 5.5 concerns the association which was uncovered between tweet-initial laughter and specific recipient-design, and some possible implications of this association. Section 5.6 contains several analyses which attempt to unpack this connection by examining instances of tweet-initial and tweet-final laughter in conversational context. First 5.6.1 examines the possibility that initial-position laughter refers back to previous or ongoing topics, and second 5.6.2 examines the possibility that initial-position laughter functions more often as an acceptance of offered laughter. Section 5.7 constitutes several gender analyses, 5.7.1 being an analysis of whether or not individual laugh forms are tweeted more often by males or by females, and 5.7.2 examining the possibility that offer and acceptance of written laughter is influenced by gender.

5.1: Overall Frequency

The unfiltered dataset which was collected represents a random sample of tweets.[10] In order to obtain an initial idea as to the relative frequencies with which different written laugh forms are used, instances of laughter in this dataset were counted and graphed in Figure 2. Several forms of laughter which were not found in the unfiltered dataset, but were found in other datasets were also included here.

Selection of forms considered to be "written laughter" for this analysis was intentionally broad. As little is known about the nature of written laughter, there was little information on for creating an informed definition of the term. Therefore this analysis considers any of the forms listed as either "emoticon" or "spelled-out" laugh forms in Section 4.2 above to be written laughter.

[10] It is unclear exactly the extent to which the sample returned by the Twitter API is *actually* random, as the precise algorithms used by the API are private, but this functionality is advertised as a random sample and has been used by numerous studies as such (Bamman et al., 2014; Weller et al., 2013).



Notice that all forms of written laughter listed in Figure 2 occur in fewer than 5% of tweets. Written laughter, while a familiar feature of discourse on Twitter, is not a feature present in a majority of tweets. 😂 and *lol* are the two most frequent laugh forms.

5.2: Co-Occurrence

Most of the analyses presented in the sections which follow will require that laugh-form datasets be reduced only to examples of isolate laughter.[11] There were, however, a relatively small number of tweets collected as part of each filtered sample which could not be used in these comparisons as they contain more than one instance of written laughter. This section is devoted to examining the frequencies with which various non-dataset laugh forms co-occur with three of the pre-specified forms of laughter (*lol*, *haha* and 😂).[12] Figure 3 displays the frequencies with which the forms

[11] Recall the definition of this term. A tweet containing *isolate laughter* contains one and only one instance of a pre-specified laugh form and contains no other laugh form as specified by the list, 😄, 😍, 😂, 😉, 😎, 😊, 😆, ☺, ☺, *lol*, *haha*, *lmao*, *lmfao*, *jk*, *hehe*, *hoho*, *ahah*, *huhu*, */chuckle*, **giggle**, and orthographic variations on these.

[12] In examples like the following:

1) @jeeessiiccaaa_: "@TheAmbitiousz1 that's too far haha the drivethru😂😂 wow okay lmao"
2) @Blake41Taylor: "@JessicaManuel95 haha so your mad because your so called number one team is going to get beat? Maybe a shout out? Haha"
3) @yes_its_james: "@pincheehacobb @CesarB_58 I am lol, no one listens to me haha"

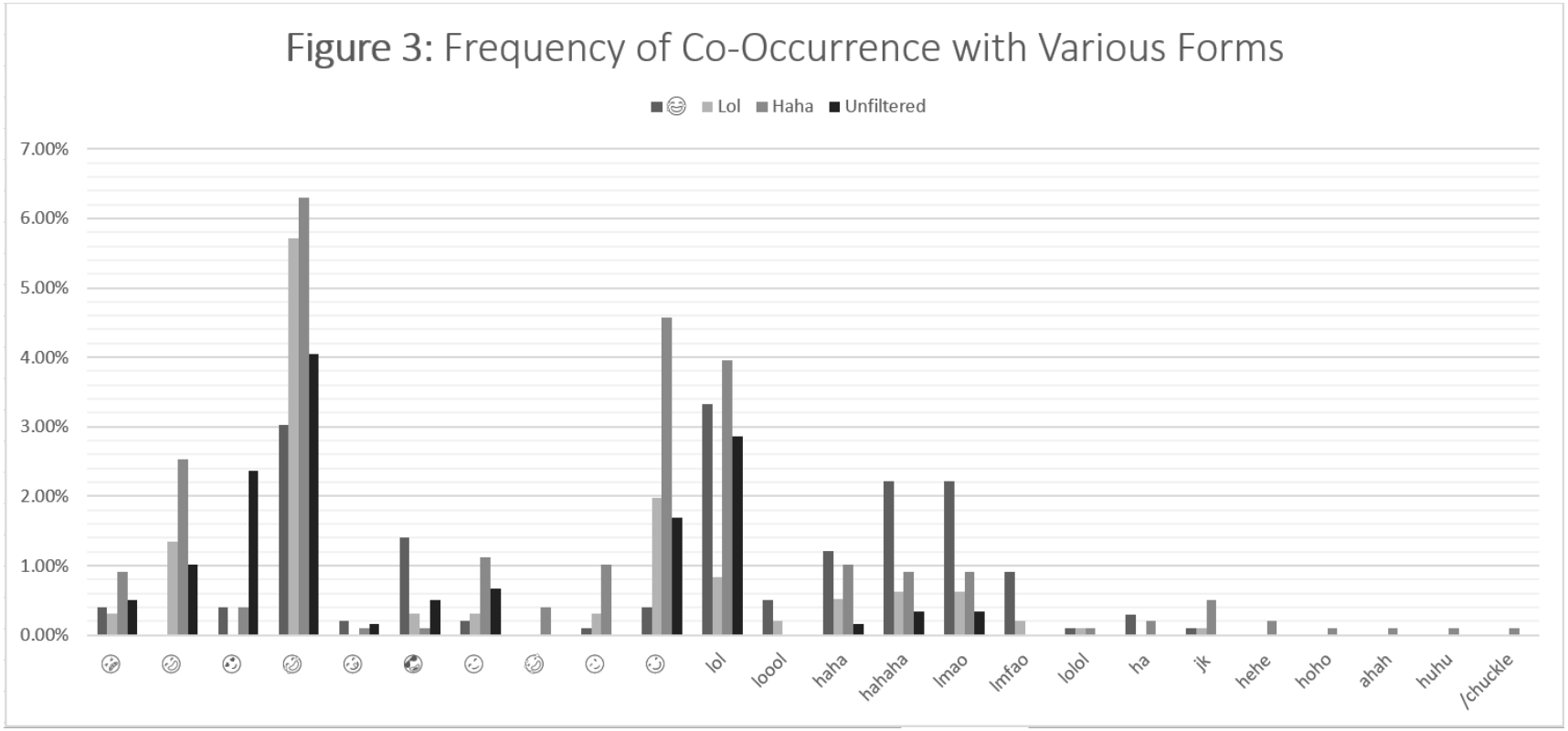


Figure 3: Frequency of Co-Occurrence with Various Forms
Lol
Haha
Unfiltered
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
lol
loool
haha
hahaha
lmao
lmfao
lolol
ha
jk
hehe
hoho
ahah
huhu
/chuckle

lol, *haha*, and 😂 co-occur with each of the forms specified for the frequency analysis in 5.1. Notice that an unfiltered column is also included. These frequencies are the same as those shown in Figure 2 above.

Haha and to a lesser extent *lol* seem to co-occur with other laugh forms, particularly emoticon laugh forms, at a higher rate than those forms would be expected to occur in an unfiltered sample. The presence of either of these two forms increases the likelihood that another laugh form will also be present. While the presence of 😂 in a tweet seems to reduce the likelihood that another emoticon will also be present, its presence does powerfully increase the probability that a spelled-out laugh form will be present.

5.3: Recipient Design

Twitter can be used either as a microblogging platform or as a short message service (SMS)[13]. Historically and canonically Twitter has been used as a microblogging platform. This means that a tweeter posts as one might to a bulletin board. Messages are public and undirected. Any user may read any tweeter's content, but readers may pick out relevant content by "following" specific tweeters. This allows them to receive only those tweeters' content on their digest page. A consequence of this is that those who post to the site are constantly conscious that their main audience consists of those individuals who have chosen to follow their accounts.

Twitter's SMS functionalities were added shortly after the site's original publication. A user desiring to use the site as a short message service will "mention" other users, typically at the beginning of the tweet, using the @username function (see Figure 1 above). Though these messages are publically accessible, so long as the tweet begins with the mention, the message will not be broadcasted automatically to all of the tweeter's followers. Instead it is sent only to the users mentioned in the

[13] Short-message Services are a broad category of internet communication mechanisms. The canonical example of an SMS is text-messaging, though any instant messaging service where messages are short and delivered (nearly) instantaneously also fall into this category.

tweet. The recipients are then notified in much the same way as they might when receiving a text message or email. Often, though not always, this targeted mentioning function is used much like a text or instant messaging service, allowing users to carry out semi-public conversations over the site.

The following examples show tweets composed using the microblog and SMS formats:

7) **Nonspecific Recipient:** "Microblog" Format:
   a. [@mitchieee47]: I hope I remember shit after this bar lmao
   b. [@_Gabby915]: Im gonna get so many toys hehe
   c. [@dabuekkehabda_x]: this has been me for the last three hours. haha makes me laugh bc it's so accurate. :)) http://t.co/fvJB2nrEWG

8) **Specific Recipient:** "SMS" Format:
   a. [@ItsOnlyZeek]: @natalietaylor96 haha you're fine that dude followed me today or yesterday I think
   b. [@Zefoh]: @SwitchSP @SPG_AX why would you leave the team to eeko lol
   c. [@Lindsay_Hecker]: @dawncflv @WriterAJCarroll @leannrimes @KarlaHoffman @kpsander @Christina4jc @danielfgalvan and it's a requirement for us to know it. lol!

These two functions provide an interesting avenue of inquiry from a conversation analytic perspective. The microblogging format allows users to publish tweets which are not designed for a specific recipient. While there is some level of recipient design, as the tweeter typically knows something about his or her pool of followers, tweets using this format are directed at a broad and flexible audience. Tweets published using the SMS format, however, are by definition specifically-targeted. Users tweeting in SMS format design their tweets to be received by a specific individual.

When evaluating the extent to which written laughter practices draw on our knowledge of spoken laughter, the distinction between specific and nonspecific recipient design is relevant for several reasons. Chiefly, it has been shown that the presence or absence of others, as well as the identities of those others, are major factors in the production or non-production of spoken laughter. More laughter is produced if others are present, especially if the individual under observation considers his or her compatriots to be friends. If written laughter is used in a way that is at all analogous to spoken laughter we would expect to see more laughter in situations in which co-participants are "more present." A user

who publishes tweets in a microblog format knows that his or her tweet will likely be read by someone in his or her pool of followers, but does not know exactly who will read it or when. The recipient is abstract. A user who publishes a tweet using the SMS format, however, knows with a reasonable level of certainty that the mentioned individual (generally someone with whom the tweeter feels comfortable conversing) will actually receive and read his or her tweet. Specifically-targeted tweets give the tweeter a greater impression that he or she is actively communicating with others. If written and spoken laughter do behave similarly, we would therefore expect to see tweets containing written laughter occur more frequently than tweets without written laughter in contexts of specific recipient design. This is, in fact, what we see.

Figure 4: Recipient Design - General

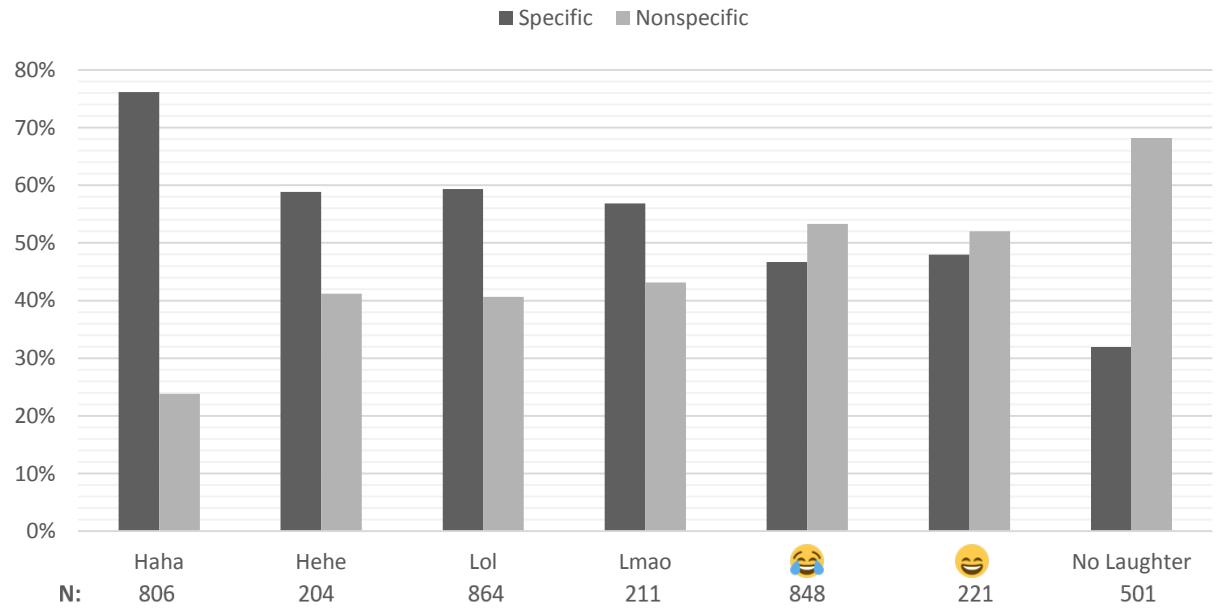


Figure 4 displays the frequency with which tweets containing each laugh form take specifically-targeted (containing @username mentions) and nonspecifically-targeted (not containing @username mentions) formats. Of the two formats, tweets containing one of the six laugh forms under observation seem to be as or more likely to be specifically-targeted as they are to be nonspecifically-targeted. This is in contrast to tweets which do not contain written laughter, which use nonspecifically-targeted format

with a frequency which is almost two times greater than the frequency with which specifically-targeted format is used.

A chi square test of independence was performed to examine the apparent relation between the presence of a laugh particle within a tweet and specific recipient design, and a relation was shown. Tweets containing no form of written laughter were significantly less likely to be designed for a specific recipient than were tweets containing written laughter $x^2(1, N=3,765)=130.37$, $p<.01$. This association between specific recipient design and the presence of written laughter, suggests that written laughter, just like spoken laughter, is used in environments in which other individuals are more "present." Written laughter is less likely to be used by tweeters in relative social isolation.

Another interesting pattern can be seen in Figure 2 as well. Tweets containing emoticon laugh forms, while more likely to be designed for specific recipients than are tweets containing no laughter whatsoever, also seem to be less likely to be specifically-targeted than are those tweets containing laugh forms which are "spelled-out" with letters (i.e. *haha*, *hehe*, *lol*, and *lmao*). A second chi square test of independence was performed in order to determine whether or not there was a relationship between the type of written laughter and the recipient design of the tweets in which the form occurs. A significant relationship was found, showing that, indeed, emoticon laugh forms are less likely to occur in specifically-designed tweets than are spelled-out laugh forms $x^2(1,N=3,154)=101.31$, $p<.01$. This difference in the recipient-design distributions for spelled-out and emoticon laugh forms implies that the two types may be used in accordance with different rules or for different purposes.

One final interesting observation which can be made in Figure 2, is that tweets containing *haha* are more likely than are tweets containing any other laugh form to be specifically-designed. These *haha* tweets are more than 15% more likely to target specific recipients than the form which is the next most-likely to be used in specifically-targeted tweets (*lol*). This may suggest that *haha* is being used in a way

which distinguishes its use both from emoticon and from the other forms of spelled-out laughter (*lol*, *lmao*, and *hehe*). This possibility will be revisited throughout this paper.

The results of this analysis are threefold. First, tweeters are more likely to target a tweet at a specific individual if it contains written laughter than if it does not. Second, tweeters are more likely to target a tweet at a specific individual if it contains a spelled-out laugh form than if it contains an emoticon laugh form. Third, tweeters seem to be more likely to target a tweet at a specific individual if it contains *haha* than if it contains any other form of written laughter under observation. Altogether three subcategories of laugh forms seem to be emerging: emoticon forms, *lol*-like forms, and *haha*. These categories seem to recur in several of the analyses below.

5.4: Laugh Locations within a Tweet

The analysis which follows examines the location of written laugh particles with regard to the boundaries of a tweet. Essentially this analysis is intended to provide general information about the distribution of written laughter within the boundaries of a conversational turn, embodied here by the boundaries of a tweet. Occasionally individuals will tweet several times in a row, but in general the end of a tweet implies the end of the speakers turn at talk.[14]

Those tweets containing isolate laughter[15] (over 90% of the total sample) were categorized as containing laughter that was either tweet-initial, tweet-medial, tweet-final, or alone. Hashtags, hyperlinks, mentions, manually retweeted passages, punctuation and emoticons were ignored when tagging a laugh form as tweet-initial or tweet-final, so long as these items were not integrated into the syntax of the message. Examples 9-12 below provide examples of tweets in each of these categories.

[14] See Sacks et al., 1974 for more information on conversational turn-taking.
[15] Recall the definition of this term, given in section 4.2. Tweets with "isolate laughter" are considered to be those tweets containing one and only one instance of written laughter where that form is the dataset-specified form.

9) **Tweet-Initial Laughter:**

> *[@BJ5995]*: @kimmy_dance lol you waiting for her to get out of a store again?

10) **Tweet-Medial Laughter:**

> *[@_kgip]*: @Dmoss_2up2down hold on lol I'll text it to you

11) **Tweet-Final Laughter:**

> *[@stackzhoe]*: I woke up like this 😩😩😩😩 party tonight tho lol #Baseboy #DumandDumber instagram.com/p/vbiRKfPseT/

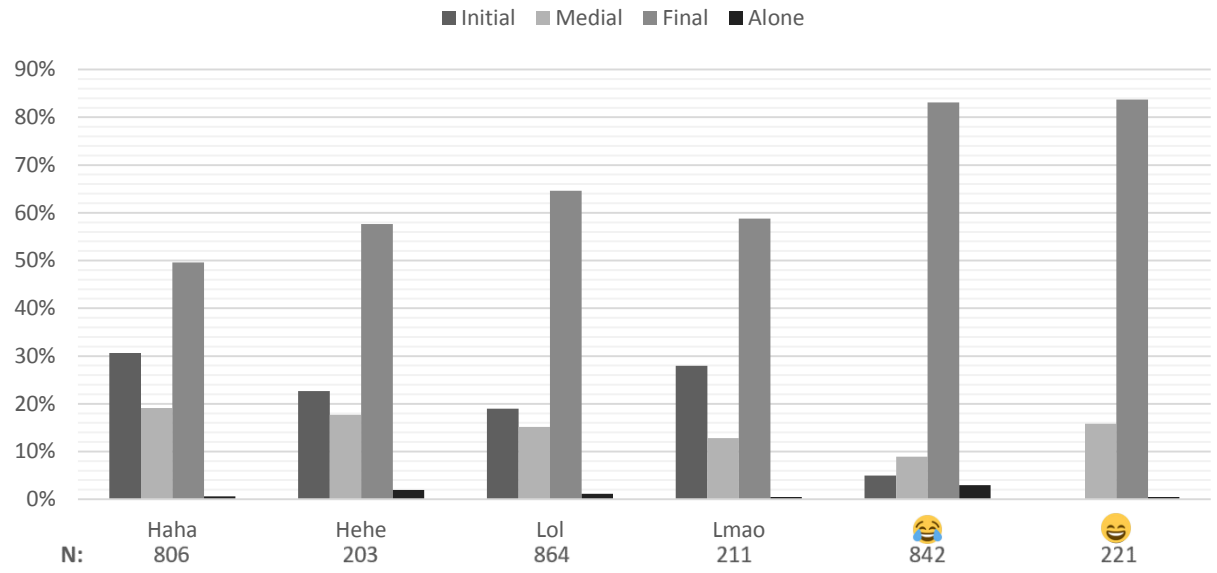
12) **Laughter Alone:**

> *[@Chappells_Show]:* "@Afroj3di: I still wanna know why I gotta shut up Joaquin" – lol

The last example above (12) contains what is known as "manual retweet." Everything between the quotation marks was originally posted by @Afroj3di, not by @Chappells_Show, the user posting here. Manual retweets are often used in the same way one might use media content (photos, links etc.): as a piece of news or as a joke. More historically this was a method by which individuals kept track of ongoing conversations (Weller et al., 2013). In order to provide a reply to a friend's tweet, a user would copy and paste the friend's tweet at the front end of her own tweet, placing it in quotation marks, and then add her own contribution at the tail end. This allows the users involved in the exchange, as well as anyone else wanting to read it, to follow the thread of a conversation. Occasionally manual retweets are still used in this fashion, and when they are I have treated quoted sections as previous turns in the conversation, not as components of the tweets in which they occur. In all cases they were ignored when logging the position of a laugh particle.

The results of this analysis can be found in Figure 5. All forms of written laughter under examination are most likely to occur at the end of a tweet and least likely to appear alone. Generally, over in over half of tweets containing any one form of laughter, that laughter occurs tweet-finally.

Figure 5: Percentage of Laugh Forms by Position within Tweet



The most interesting feature of the analysis in Figure 5 is that the emoticon laugh forms (😄 and 😂) almost never appear in tweet-initial position. While the spelled-out forms are most likely to occur tweet-finally, they do appear tweet-initially at substantial rates. This observation provides more evidence for the proposal made above that spelled-out and emoticon forms of laughter may be used according to different models. Though *haha* is the most likely of the spelled-out forms to be used initially, the difference here is negligible. Thus, this analysis shows no real difference between the use of *haha* and the use of any other spelled-out laugh form.

It is worth noting that each form listed above occurs non-medially over 80% of the time. Written laughter, in general, tends to bookend tweets. It has been observed that emoticons tend to appear only at the ends of phrases, much like phrase-final punctuation (Provine et al., 2007). A cursory overview of the spelled-out laughter datasets collected for this study suggests that laughter, unlike emoticons, can be used at the beginnings of phrases as well. As tweets are limited in length, it may be the case that the fact that written laugh forms bookend *tweets* is a side-effect of the fact that these

laugh forms bookend *phrases*. While the positioning of written laughter at a sentence level would surely be a productive route for future inquiry, this is an analysis of written laughter at a tweet level, and so a sentence-level analysis is left for future work.

5.5: Recipient Design and Laugh Location

The two comparisons above showed that, while there are some similarities between how spelled-out laugh forms and emoticon laugh forms behave with regard to recipient-design distributions and placement within a tweet, there are also some differences. Though tweets containing any form of written laughter are more likely to be specifically targeted than tweets which do not contain laughter, tweets containing the spelled-out laugh forms *lol*, *haha*, *hehe*, and *lmao* are even more likely to be designed for a specific recipient, and are more likely to contain tweet-initial laughter than are tweets which use the emoticon forms 😂 and 😄. This section is devoted to the following question: Are these two observations, that spelled-out forms are more likely to occur tweet-initially than are emoticon laugh forms and that tweets containing spelled-out laugh forms are more likely to be specifically-targeted than tweets containing emoticon laugh forms, related? A hypothesized connection between the frequency with which a laugh form appears initially and the frequency with which it occurs in specifically-targeted environments is appealing chiefly because initial laughter is strongly correlated with specific recipient design.

Figure 6 shows the recipient design patterns for tweets containing isolate instances of various tweet-initial laugh forms.[16] The column for the 😊 form is empty because the collected dataset contained no isolate examples of initial 😊.

For all five laugh forms for which information about initial isolate laughter is available, initial isolate laughter occurs much more often in tweets with specific recipient design. Though the one

[16] See example 9 above.

emoticon form that occurs initially, the laughter with tears emoticon, does show an association between tweet-initial 😂 and specific recipient design, it occurs about 15% more often in nonspecifically-targeted tweets than do the spelled-out initial forms.

Figure 6: Recipient Design - Tweet-Initial Laughter

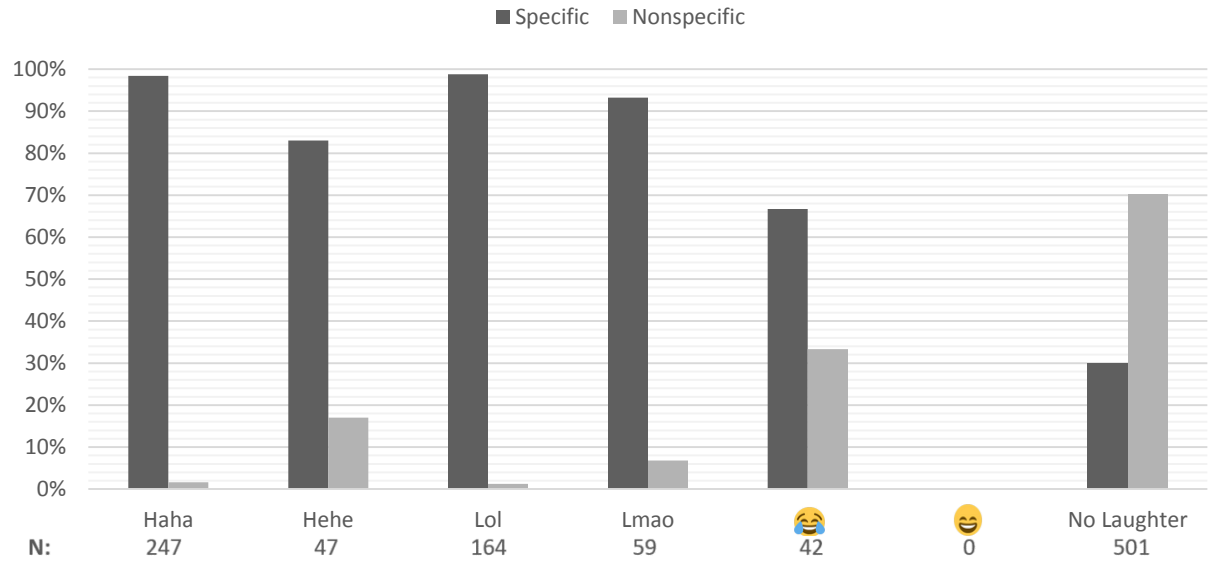
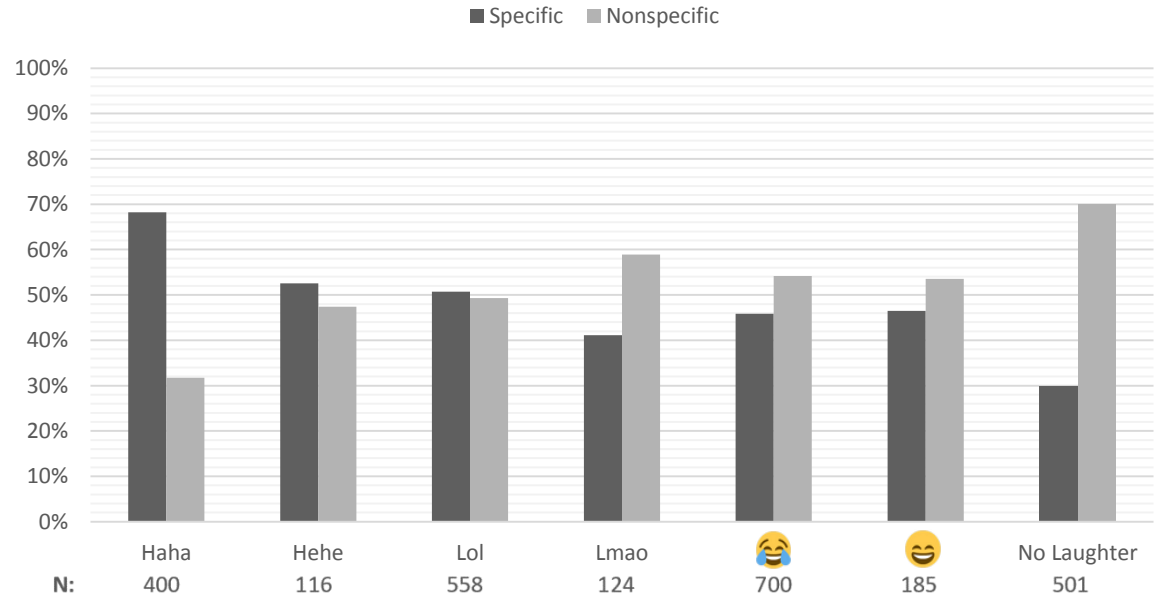


Figure 7: Recipient Design - Tweet-Final Laughter



Meanwhile, Figure 7 shows the recipient-design breakdowns for those tweets containing tweet-final isolate laughter.[17] The connection between tweet-final laughter and specific recipient design seems far more tenuous. Though, for all six laugh forms, tweets containing final laughter are more likely to receive specific targeting than are tweets which do not contain laughter, the behavioral differences between initial and final written laughter are quite noticeable. In tweet-final position, the spelled-out laugh forms *hehe*, *lol*, and *lmao* show recipient-design frequencies closely resembling those shown by the emoticon forms. The one form for which tweet-final laughter is still clearly associated with specific targeting is *haha*, and in this its behavior is set apart from the other five written laugh forms.

Spelled-out laugh forms are used more often initially than are emoticon laugh forms. Initial laughter is used almost exclusively in specifically-targeted tweets. Might the reason that certain laugh forms have high rates specific-targeting be that these same laugh forms are used more often in the strongly-specific tweet-initial position? Can we legitimately consider the rate with which a form receives specific recipient design to be dependent on the rate with which that form is used tweet-initially?

The charts in Figures 8 and 9 below represent comparisons between the rate with which each form occurs tweet-initially (see 5.4) and the rate with which tweets containing each form receive specific recipient design (see 5.3). Figure 8 examines the possibility that there is a linear relationship between these two variables, namely those laugh forms which occur more often initially occur more often in specifically-targeted environments. Figure 9 instead depicts the suggestion that has been made several times above, that the various laugh forms under observation are used in accordance with three separate general models: emoticon laughter, *lol*-like laughter (*lol*, *lmao*, *hehe*), and *haha*.

The suggested linear model drawn in Figure 8 serves to account fairly well for four of the forms, 😂, 😊, *lol*, and *hehe*, but fails to adequately account for *haha* and *lmao*. There does, however, appear to be a generally positive trend, making this model intriguing. The addition of more data points through

[17] See example 11 above

the analysis of more laugh forms might show, in the future, that this way of considering the data is more generally powerful than it appears to be here. However, given the fact that this study has made only six data points available, the model proposed in Figure 9 is currently more appropriate.

Figure 8: Specific-Targeting and Initialness - Proposed Linear Relationship

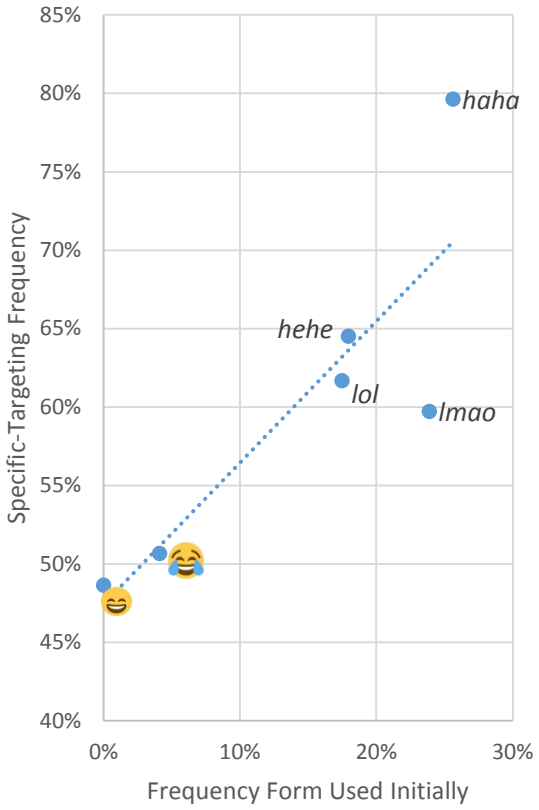
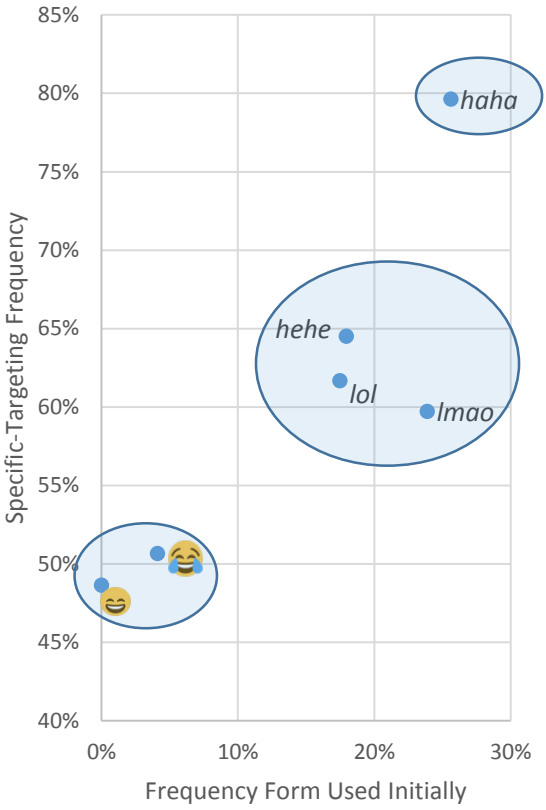


Figure 9: Specific-Targeting and Initialness - Three Proposed Patterns

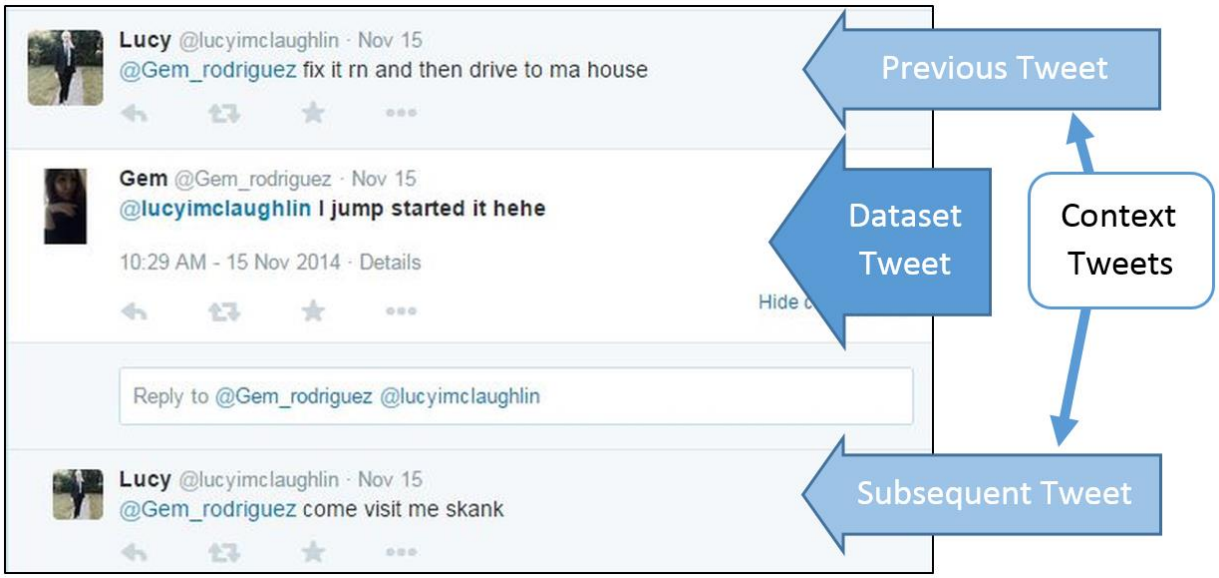


It is of note that at no point in this analysis thus far has any real difference in the use of abbreviated laugh forms (*lol*, *lmao*) and onomatopoeic laugh forms (*haha*, *hehe*) been evident. Though *haha* seems to show some differences from the other three, the usage patterns of *hehe* are very similar to those of *lol*. The three groups which seem to be emerging are, rather, those shown in Figure 9.

5.6: Location of Laughter in Conversational Context

The sections above revealed intriguing differences between the recipient design distributions of tweets containing tweet-initial laughter and tweets containing tweet-final laughter. In this section tweet-initial and tweet-final laughter are compared with regard to the nature of conversationally adjacent tweets.

Figure 10: A Dataset Tweet and its Context Tweets



Unlike the sections above, this section, as well as parts of Section 5.7, discuss the relationships between tweets which are part of the gathered datasets and tweets in their immediate conversational context. Take for an example a tweet, such as the one in Figure 10, that was collected as a part of the dataset of tweets containing *hehe*. If, the user composing this collected *dataset tweet* did so in order to respond to something tweeted earlier by another user, this *previous tweet* was collected. If, sometime after collection,[18] another user posted a response to the collected dataset tweet, this *subsequent tweet* was collected as well. These previous and subsequent tweets, which will be referred to collectively as *context tweets* below, provide information about the dataset tweet in the same way that the other

[18] Recall that because data collection took place in real time conversational context was collected some time later in order to give any users who might choose to respond to dataset tweets the time to do so.

forms of tagged metadata do. Just as we know that the tweet in Figure 10 contains tweet-final *hehe*, we know that it has a previous and subsequent tweet, both tweeted by a female,[19] neither of which contains written laughter. When in the sections that follow I refer to a dataset's previous or subsequent tweets, I am referring to any previous tweets or any subsequent tweets which were logged as occurring before or after any of that dataset's dataset tweets.

Recall that these context tweets were collected using the "view conversation" function on Twitter's website.[20] Unfortunately, the algorithm which governs this functionality is imperfect. This means that for some dataset tweets which *did* respond to previous tweets or elicit subsequent tweets, these context tweets could not be retrieved. The algorithm seemed to miss subsequent tweets more often than it did previous tweets. In addition to this, there were rare circumstances in which those users who composed the context tweets in question had altered their privacy settings so as to make it impossible for anyone without specific permissions to view their posted content. These two issues with the retrieval of context tweets, while they do not invalidate the information provided by those context tweets which *were* collected, does invalidate any statistic which examines, for example, the frequency with which tweets in the *hehe* dataset have previous tweets or subsequent tweets.

Therefore it is *vital* to remember that when, in the sections that follow, I refer to the frequency with which the context tweets of a given dataset possess a trait, I am *not* referring to the frequency with which dataset tweets follow or precede tweets with that trait. Instead I am referring to the frequency with which those context tweets that *were* collected for a given dataset have that trait. For example, when, in section 5.6.2, I refer to the frequency with which the previous tweets of the *haha* dataset contain spelled-out laughter, I am referring to the percentage of the previous tweets *collected* for the

[19] For gender-identification techniques, see Section 5.7
[20] See section 4.2

haha dataset contained spelled-out laughter, not the percentage of tweets containing *haha* that were preceded by tweets containing spelled-out laughter.

Two major analyses were conducted concerning differences in the conversational purpose of tweet-initial written laughter and tweet-final written laughter. The first, in section 5.6.1, considers the possibility that tweet-initial laughter refers back to a previous or ongoing topic, while tweet-final laughter more often refers to tweet-internal topics. The second analysis, outlined in 5.6.2, considers parallels between the placement of spoken laughter within a conversational turn, and the placement of written laugh particles within a tweet. It is in this section that the offer/acceptance patterning of face-to-face laughter is investigated with regard to written laughter.

5.6.1: Tweet-Initial vs. Tweet-Final Laughter: Externally-Referent Pronouns

This analysis considers the possibility that tweet-initial laughter tends to reference an ongoing topic or laughable first mentioned in a previous conversational turn, while final laughter tends to reference a laughable produced by the tweeter. Studies of face-to-face laughter have often suggested that turn-initial laughter tends to have utterance-external referents, as opposed to turn-final laughter which often references the speaker's own utterance (Glenn, 2003; Holt, 2011). If this were the case for written laughter as well, it could go a long way to explaining the strong association between tweet-initial laughter and specific recipient design, as a previous or ongoing topic must be described in a previous turn, something which can only exist in a specifically-targeted exchange. It would also still allow for the relatively high, though less-pronounced frequency of specific recipient design among instances of tweet-final laughter. Some examples of exchanges for which initial laughter seems to reference an ongoing topic can be found in Appendix 2.

The most obvious way to approach this question would be to examine the frequency with which dataset tweets containing tweet-initial laughter have preceding tweets and compare that to the

frequency with which dataset tweets containing tweet-final laughs have the same. Unfortunately, for the reasons described in the introduction of this section, this comparison could not be made.

Instead, another method was used to determine whether tweet-initial and tweet-final laughter differ in their ability to reference an external topic. This analysis examines the presence of externally-referent pronouns in the "punctuation sentence" containing the laugh particle. The relevance of this analysis is based on a single assumption. This assumption is that where written laugh particles occur in close proximity to externally-referent pronouns, they are more likely to also refer to a tweet-external, previous or ongoing topic.

For the purposes of this study, an externally-referent pronoun is defined as a third-person pronoun with a meaning that cannot be inferred from the text of the tweet alone. It refers back to a topic discussed earlier in a conversation. The tweet in [13], for instance, contains an explicit externally referent pronoun (*it*). Notice that a reader reading [13] in isolation does not know what "it" means.

13) **Externally-Referent Pronoun (Explicit):**

> [*@Knee_Uhhhhh*]: @K_Soto214 **it's** inactive lol.

This is in contrast to examples like [14] below, in which the pronoun "him" is clearly co-referenced with "that cute guy in the hat."

14) **Internally-Referent Pronoun:**

> [*@MetalEmpress*]: @xHollyGlambertx All I can think about when I see this photo is Adam pushing that cute guy in the hat against a wall & making out w/ **him** lol

There is a second category of externally-referent pronouns which has been included in this analysis. These are the implied externally-referent pronouns. Notice that in [15] there is an implied "it" or "that" before "not gonna happen."

15) **Externally-Referent Pronoun (Implicit):**

> [*@jduccee*]: @JennPortilloo not gonna happen lol

The presence or absence of an externally-referent pronoun was evaluated within the "punctuation sentence" containing the laugh particle in question. This is a loose definition merely intended to restrict the portion of a tweet under scrutiny to the general area of the tweet in which the laughter actually occurs, particularly in cases where tweets contain more than one sentence. The borders of the punctuation sentence were marked wherever traditional sentence-separating punctuation, emoticons, hyperlinks, or hashtags were introduced, as well as at the beginnings and ends of a tweet.

Figure 11: Initial *lol* refers to an earlier subject without an externally-referent pronoun

It is important to note that while the presence of an externally-referent pronoun may imply a topic-continuation, the pronoun need not be present in order for that tweet to *be* a topic continuation. Figure 11 below provides an example of an instance in which *lol* in the dataset tweet is clearly being used to refer back to the comment made in the previous tweet, but an externally-referent pronoun is not used. In the datasets collected for this project there are many examples of tweets which seem to bear some sort of external reference, but leave it unclear whether the laugh particle references the previous topic or the specific contribution being made by the current tweeter.[21] Such tweets were

[21] To provide a few examples:

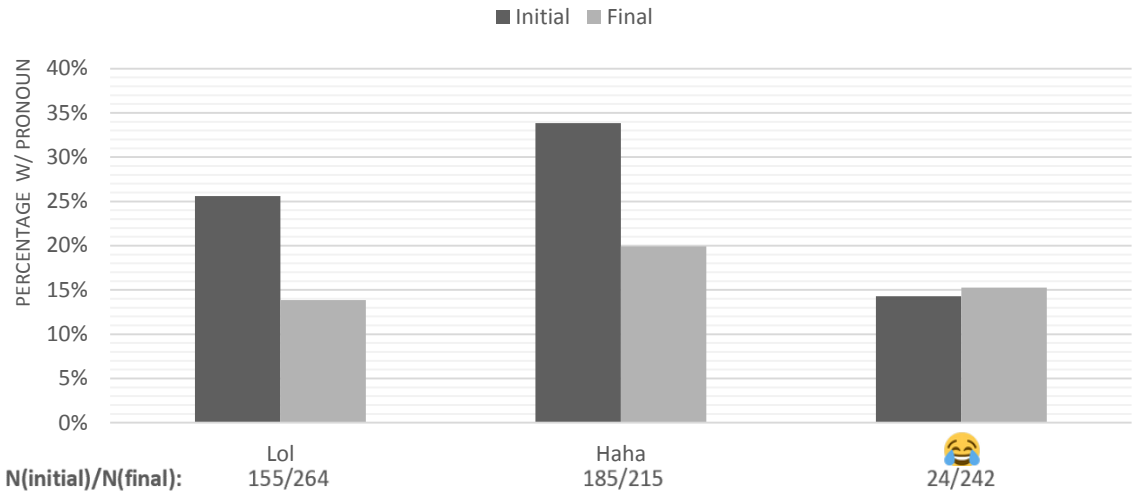
1. [@Blake41Taylor]: @JessicaManuel95 haha so your mad because your so called number one team is going to get beat?
2. [@whit_law]: *@Jro_Chicago haha I'll be keeping up with the game, tony has it on his phone!*
3. [@_azucenaamonic]: *@sayhelloirene @559_emm lol no life*

extremely difficult to tag in a reliable and consistent manner. The choice to tag only those instances of external references that make use of externally-referent pronouns served to provide a far less ambiguous set of tweets in which external reference was clear. Unfortunately, this meant that only the absolute clearest cases of external reference were considered as such for this analysis. The actual rates of external reference were likely quite a bit higher, both for tweets containing tweet-initial laughter and for tweets containing tweet-final laughter.

As tagging the presence of externally-referent pronouns must be done by hand, and requires large datasets in order to provide powerful results, this analysis was conducted using only the three large-sample datasets: *lol*, *haha*, and 😂. These datasets were originally enlarged because they exemplify both the three initialness/specificity categories drawn up in Figure 9 above (emoticon, *lol*-like, and *haha*) and the three major categories of netspeak items often discussed in other studies (emoticon, abbreviation, and onomatopoeia). The presence or absence of externally-referent pronouns within the same punctuation sentence was marked for tweets containing tweet-initial isolate laughter and containing tweet-final isolate laughter. The results of this comparison are shown in Figure 12. Percentages are given in terms of tweets with that laughter form *in that position*, rather than simply in terms of the number tweets *containing* that laughter form.

If we assume that tweets containing tweet-initial laughter do, in fact, tend to reference a previous or ongoing topic then we would expect to see tweets containing tweet-initial written laughter to contain externally-referent pronouns at a higher rate than do tweets containing tweet-final written laughter. As far as *haha* and *lol* are concerned, this is, in fact, what we see in Figure 12. Contrastively, in the case of 😂, this trend cannot be shown. However, since the number of examples of tweet-initial 😂 was severely limited, this result may be deceiving.

Figure 12: Percentage of Tweets Containing Three Laugh Forms in Tweet-Initial and Tweet-Final Position which Contain Externally-Referent Pronouns



Though these statistics are drawn from relatively small samples, it may prove worthy of note that for all three forms the frequency with which externally-referent pronouns co-occur with tweet-final laughter seems similar (hovering between 13 and 20%). Incidentally, this frequency is the same as the apparent frequency with which tweet-initial 😂 co-occurs with externally-referent pronouns (though this frequency was drawn from a small dataset, N=24). If these statistics are representative, this analysis would suggest not only that tweet-initial and tweet-final 😂 behave similarly to one another, but that both behave similarly to tweet-final *lol* and *haha*. Perhaps emoticon forms are somehow inherently internally-referent; referring only to the specific contribution being made by an individual tweeter. This is approximately what has been proposed by those who discuss emoticons as indicators of illocutionary force (Dresner & Herring, 2010). They argue that emoticons provide information about intended deeper meaning of the writer's own contributions to the conversation.

Perhaps spelled-out laugh forms like *lol* and *haha* can use initial position to indicate that the laughable is external. If we then suppose that an emoticon cannot take an external referent, the

fronting of an emoticon would not have the same power as the fronting of a spelled-out laugh form. This would explain the lack of distinction between tweet-initial and tweet-final laughter in Figure 12, but it might also explain why emoticons appear relatively rarely tweet-initially.

While *haha* shows the association between tweet-initial position and the presence of externally referent pronouns, it also shows generally higher rates of co-occurrence with external reference. This may explain the high rates of specific recipient design seen earlier. Where 😂 is strictly internally-referent, *haha* may be somewhat *externally*-referent. If *haha* tends to refer to previous or ongoing topics that might explain why it is so often found in environments that allow for the presence of an ongoing or previous topic, environments of specific recipient design. This association could, of course, be drawn in the opposite direction. If a form is more often used in an interpersonal, specifically-targeted environment, it would make sense for that form to more often reference the surrounding conversational context as well.

5.6.2: Presence/Absence of Conversationally-Adjacent Laughter

Many of the studies discussed in the literature review of this paper have found that when laughter is used in face-to-face conversation it is both offered in accepted. One speaker will laugh towards the end of her utterance, offering laughter, and her conversational partner will begin his utterance by laughing as well, accepting her offered laughter (Glenn, 1989, 1991, 2003; Jefferson et al., 1977; Provine, 2001). This section examines the possibility that written laughter may be used in a similar manner. The comparisons which follow look for associations between presence and/or location of certain forms of written laughter in a dataset tweet and the presence of laughter in its context tweets.

Ideally this analysis would include a comparison for which the locations of laugh forms in dataset tweets *and* in context were variables. Unfortunately, the data gathered for this project did not

provide sufficient sample sizes to conduct such analysis productively and therefore this aspect of the study is left for future work.

In the analyses which follow, context tweets will be tagged as containing or not containing spelled-out and emoticon laughter. The distinction made here between emoticon laugh forms and spelled-out laugh forms was motivated by those of the results presented above which seem to suggest that emoticon forms are used on Twitter according to a different model than are spelled-out forms. Most of the analyses in this section concern the presence or absence of spelled-out forms in context tweets, but a mention of emoticon laughter is made towards the end of this section.

The datasets involved here were filtered down once again to only examples of isolate laughter, as the location of the laugh particle in a dataset tweet was a major factor in this analysis. Each dataset tweet's context tweets were tagged if they contained spelled-out or emoticon laughter, using the lists of laugh forms given in 4.2.[22] This means that the factor being tested in each case was the influence each dataset laugh form had on the presence or absence of previous laughter of these two types in general, not the influence these forms had on the specific identity of these forms.

Throughout this section I will refer to the frequency with which the context tweets for a certain group of dataset tweets (such as those tweets containing tweet-initial *haha*) contain laughter. For the reasons given in the introduction of 5.6, these frequencies will be the percentage of the context tweets collected for that group of dataset tweets which contain laughter. This means that if I am defining the frequency with which those tweets containing tweet-initial *haha* have previous tweets containing laughter, I am reporting the frequency with which the previous tweets of tweet-initial *haha* contain laughter, not the frequency with which tweets containing *haha* both contain tweet-initial *haha* and are preceded by a laughing tweet.

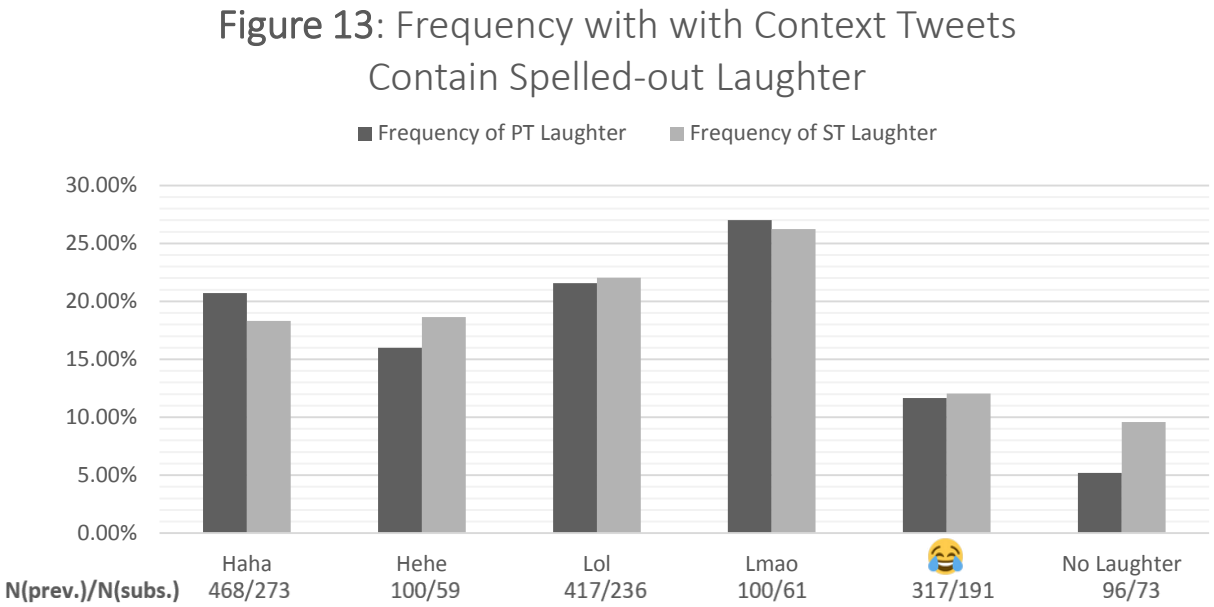
[22] Spelled-out forms: *lol*, *haha*, *lmao*, *lmfao*, *jk*, *hehe*, *hoho*, *ahaha*, *huhu*, **giggles**, */chuckle* and orthographic variations on these. Emoticon forms: 😄, 😍, 😂, 😉, 😎, ☺, 😊, ☺, ☺, and their ASCII counterparts (any form with an upturned mouth was included).

This analysis examines data from those of the unfiltered tweets which contain no laughter, as well as data from five out of the six collected datasets of laugh forms. The one form not discussed in this section (😊) has been left out because no examples were found in which this form was used initially in isolate. The location of the dataset laugh form within a tweet is a major variable in the analyses which are to follow, and so 😊 has been put aside for the time being.

5.6.2.1: General Association between the Presence of Various Written Laugh Forms and the Presence of Spelled-Out Laughter in Context Tweets

This first analysis examines the relationship between the presence of written laughter within a dataset tweet and the presence of spelled-out laughter in its context tweets without regard to the location of the dataset form within the tweet. This is intended to establish first, whether or not the presence of the dataset forms selected is correlated with the presence of spelled-out laughter in context tweets in general, and second, whether any of these laugh forms is preferentially used before or after other tweets containing written laughter. If, for example, one laugh form had been shown to be associated only with the presence of spelled-out laughter in its previous tweets and not its subsequent tweets, this laugh form may have served a more responsive purpose.

The prediction of this analysis is that the presence of a written laughter in a dataset tweet is associated with the presence of spelled-out laughter in its context tweets. This would lead us to predict that the frequency with which the context tweets of dataset tweets containing each written laugh form contain spelled-out laughter ought to be higher than the frequency with which the context tweets of dataset tweets containing no written laughter do the same. This prediction has two halves. The first half predicts that each form will be associated with elevated frequencies of spelled-out laughter within its previous tweets, while the second predicts that each form will be associated with elevated frequencies of spelled-out laughter within its subsequent tweets. Several of the forms in question show both predicted associations, as can be seen in Figure 13



Notice that dataset tweets containing all four of the spelled-out laugh forms under observation seem more likely to have context tweets, both previous and subsequent, that contain spelled-out laughter than are the dataset tweets containing no laughter. This is exactly the result that was predicted. However, though dataset tweets containing 😂 are slightly more likely to have context tweets which contain spelled-out laughter than are dataset tweets containing no laughter, this difference is very small. This seems to be no strong association between the presence of 😂 in a tweet and the presence of spelled-out laughter in either its previous or subsequent tweet. This provides further evidence that emoticon and spelled-out forms of written laughter are used differently. This particular difference will be revisited in section 5.6.2.3 below.

Another observation that can be made based on the information given in Figure 13 is that for each form, dataset tweets are about as likely to contain spelled-out laughter in previous tweets as they are to contain spelled-out laughter in subsequent tweets. There is no form which is primarily used before tweets containing laughter and no one form which is primarily used after tweets containing laughter.

5.6.2.2: Adding Laugh Location as a Factor

This section examines the possibility that the location of a laugh form within a tweet may have an influence on the frequency with which its previous tweets and subsequent tweets each contain laughter. This question stems from the observation, made in studies of the offer/acceptance patterns for face-to-face laughter that offered laughter tends to occur towards the end of an utterance while accepted laughter tends to occur towards the beginning of an utterance. Examples of the various combinations of previous and subsequent laughter and locations of laugh particles within tweets can be found in Appendix 3.

If written laughter follows similar offer/acceptance patterns to face-to-face laughter, we might expect to see initial laughter more strongly associated with laughter in a previous tweet than is final laughter, and final laughter more strongly associated with laughter in a subsequent tweet than is initial laughter. We are essentially looking for two separate predictions here. The first prediction, that tweet-initial laughter is more strongly associated with laughter in a previous tweet than is final laughter, we will call the "Initial-Previous Relation." The second prediction, that tweet-final laughter is more strongly associated with laughter in a subsequent tweet than is initial laughter, we will call the "Final-Subsequent Relation"

Figures 14 and 15 show that Initial-Previous and Final-Subsequent Relations are observable at least for some laugh forms. The Initial-Previous Relation, shown in Figure 12, is observable for *hehe*, *lmao*, *lol*, and 😂, but, interestingly, not for *haha*. Meanwhile, the Final-Subsequent Relation, shown in Figure 13, is observable for *haha*, *hehe*, *lol*, and *lmao*, though the association only seems strong for *hehe* and *lmao*.

Figure 14: Frequency of Spelled-Out Laughter in Previous Tweets as a Function of Dataset Laugh Location

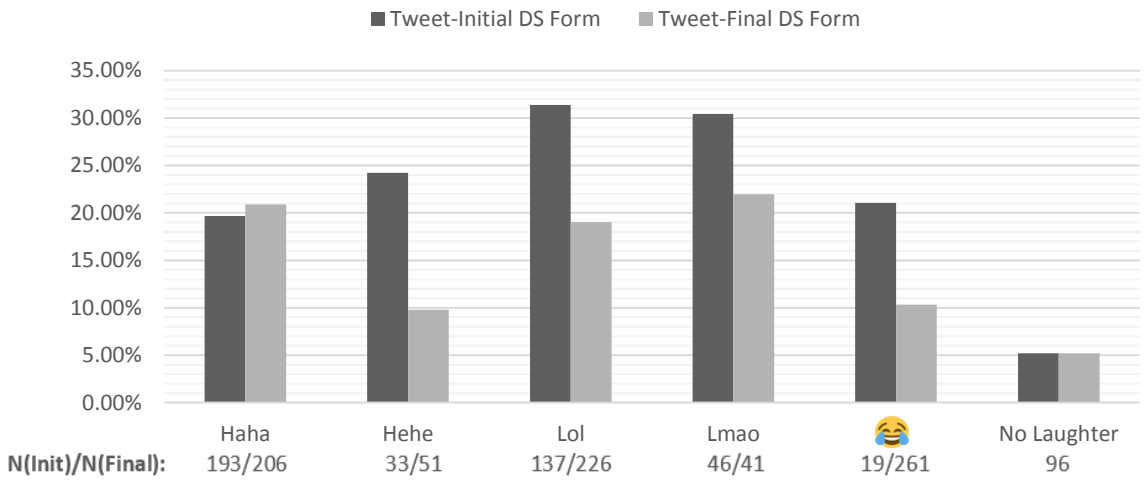
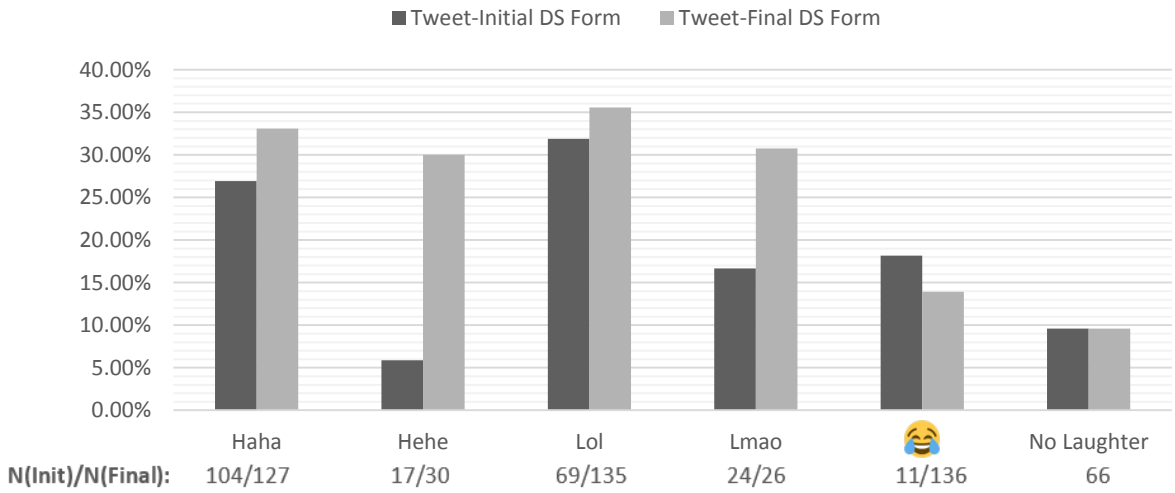


Figure 15: Frequency of Spelled-Out Laughter in Subsequent Tweets as a Function of Dataset Laugh Location



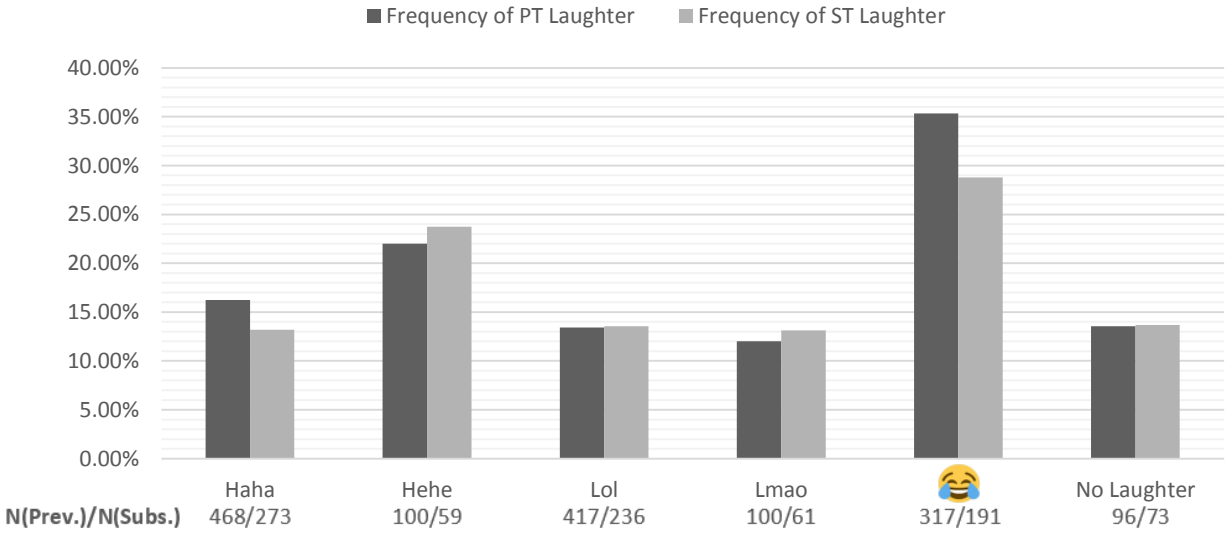
Interestingly, the Initial-Previous Relation seems more consistent than the Final-Subsequent one. This may be partially due to the fact that the number of subsequent tweets collected for each dataset was smaller than the number of previous tweets collected, or it may be because tweet-initial

laughter behaves more systematically as a laughter acceptance than tweet-final laughter does as an invitation. Either way, some forms do show an association between tweet-initial written laughter and the presence of spelled out laughter in a previous tweet, and some forms do show an association between tweet-final written laughter and the presence of spelled-out laughter in a subsequent tweet.

5.6.2.3: General Association between the Presence of Various Written Laugh Forms and the Presence of Emoticon Laughter in Context Tweets

The above two sections have examined the relationship between the presence of various dataset laugh forms and the frequency with which context tweets contain *spelled-out* laughter. This section looks at the relationship between the presence of these dataset laugh forms and the frequency with which context tweets contain *emoticon* laughter. Here we go back to a prediction similar to the one made in 5.6.2.1, that those dataset tweets containing some form of written laughter will be more likely to have context tweets which contain some form of emoticon laughter than are dataset tweets containing no laughter whatsoever.

Figure 18: Frequency with which Context Tweets Contain Emoticon Laughter



In examining the results of this comparison, shown in Figure 18, it becomes immediately obvious that dataset tweets containing 😂 seem to be more likely than dataset tweets containing any other form of written laughter under observation to have context tweets which contain emoticon laughter. This is the opposite of what we saw in 5.6.2.1, where 😂 was the only form which was not associated with spelled-out laughter in previous and subsequent tweets.

A second observation which can be made about Figure 18 is that the presence of at least three of the spelled-out forms, *haha*, *lol*, and *lmao*, does not appear to be associated at all with the presence of emoticon laughter in context tweets. The one spelled-out laugh form which does show the predicted association, *hehe*, shows it only weakly. Yet again a difference between the spelled-out laughter forms and emoticon forms is evident.

It must be noted that for this comparison tagging of emoticon laughter was, perhaps, overly flexible. If a large number of emoticon forms which may or may not have actually represented laughter were tagged as such. If emoticon forms were in fact over-tagged, they may have created enough noise to have obscured associations which may otherwise have been present, or to have created associations which otherwise would have been absent.

5.7: Gender Analyses

The following sections represent two analyses concerning the effect of gender on the use of written laughter. We have seen above that some of the offer/acceptance patterns seen in the use of face-to-face laughter can be observed in the conversational use of written laughter as well. This section is motivated by two observations about laughter and gender made fairly often in the literature. First, many of the studies of netspeak languages have found that most “emotive” expressions such as the laugh forms under observation in this thesis are used much more often by females than by males. Second, many studies of face-to-face laughter have found that the extent to which offered laughter is

accepted or not accepted is strongly dependent on the gender of the participants. The first part of this section delves into the first observation, looking specifically at whether the forms under observation tend to be used more by females or by males in these datasets. The second part of this section concerns the second observation, that the likelihood that an individual will accept offered laughter is largely dependent on the gender of the individuals involved. This second part examines the influence of gender on offer/acceptance patterns of written laughter.

Twitter, unlike many other social networking platforms, neither requires nor asks users to attach to their account the gender with which they identify. In fact, all information attached to any individual's account may be completely fabricated. A user may tweet as Justin Bieber, Mockingjay❤, Batman, or even God. Most studies examining gendered use of language on Twitter have therefore followed the accounts of subjects whose gender identity had been previously ascertained.

As data for this study were collected as a random sample, rather than by following specific individuals whose gender identity was known no real-world information about each tweeter was known. One previous study (Bamman et al., 2014), did identify the gender of anonymous Twitter users by running users' volunteered first names through a computer program intended to identify names as male or female. If a first name could be found at least 1000 times in the most recent available US Census and was over 80% associated with individuals of one gender or the other (names like Paula, James, or Desiree), that users' data was tagged as either male or female and was used for their project.

In a pilot study I conducted, this method was applied to the *lol* dataset of this project. The methods above allowed only about 1/3 of users to be identified as male or female. Even when tagging was done by hand, allowing nontraditional spellings (ex: *MAKAiLA*) or gendered inventions (ex: *NAWTY❤GAL*), only about 55% of the tweeters in the dataset could be identified as male or female based on username alone. For a small, but not negligible portion of these identifications other

information on the tweeter's homepage suggested that the gender categorization based on username alone had, in fact, yielded incorrect results.

A second pilot study was therefore conducted to search for a better identification method. I examined the profile pictures associated with individuals in two friends' Twitter networks. Both users were aware of the real-world gender identification of most of the users in their networks so my identifications could be compared with their expert knowledge. Members of these two twitter networks were identified as male or female based on the gender of the person or fictional character in their profile pictures. Groups photos containing only members of one gender were also accepted, and photographs of celebrities were not included. Not every person in these two Twitter networks could be identified using this method, but in general a larger percentage of users could be identified using this method than the previous methods attempted, and these identifications seemed to be more consistently accurate than previous methods. I could identify the gender of 73.78% of the first individual's network and less than 0.05% of those identifications were inaccurate. In the second individual's network, tagging was less successful. 58.72% of this individual's total network could be gender-identified, though as 24.25% if his twitter network was ungendered businesses and blogs, this rate is not surprising. Of those pages followed by this second individual that *were* gendered 77.52% could be identified. Only 2.90% of the total identifications were explicitly incorrect, though an additional 8.70% attached genders to non-gendered entities such as blogs and news agencies. Though this method of identification was much more labor-intensive and was still not completely accurate, it was a method better suited to this project.

An initial attempt to identify users in the dataset for this project using their profile pictures revealed that about 80% of users could be identified as male or female if single-person gendered photographs, or single-character gendered non-photographs were used as an identification method. Those individuals who could not be identified had profile pictures showing a character of ambiguous

gender, a photograph containing individuals of multiple genders, photographs for which gender could not be established, or pictures which contained neither characters nor photographs of people (such as an advertisement for an event, or a photo of a sunset).

This second method was the method used for gender identification in this study. For each tweet in the dataset the gender of the tweeter as well as the genders of any other tweeters found in the conversational context of the dataset tweet were marked. Particular focus was given to the dataset tweeter and the user who tweeted directly before and directly after that individual. In the end identification rates were only slightly lower than the 80% identifiable in the pilot (hovering around 75%, depending on the laugh form in question.) The precise percentage of each dataset for which gender could be identified is given in Table 3. In the analyses which follow only that portion of the data for which gender could be identified was used.

Table 10: Identification Percentages

Form	Identified	Total	% Identified
Haha	782	1000	78.20%
Hehe	179	250	71.60%
Lol	776	1000	77.60%
Lmao	200	250	80.00%
😂	790	1000	79.00%
😄	170	250	68.00%
Unfiltered	428	593	72.18%

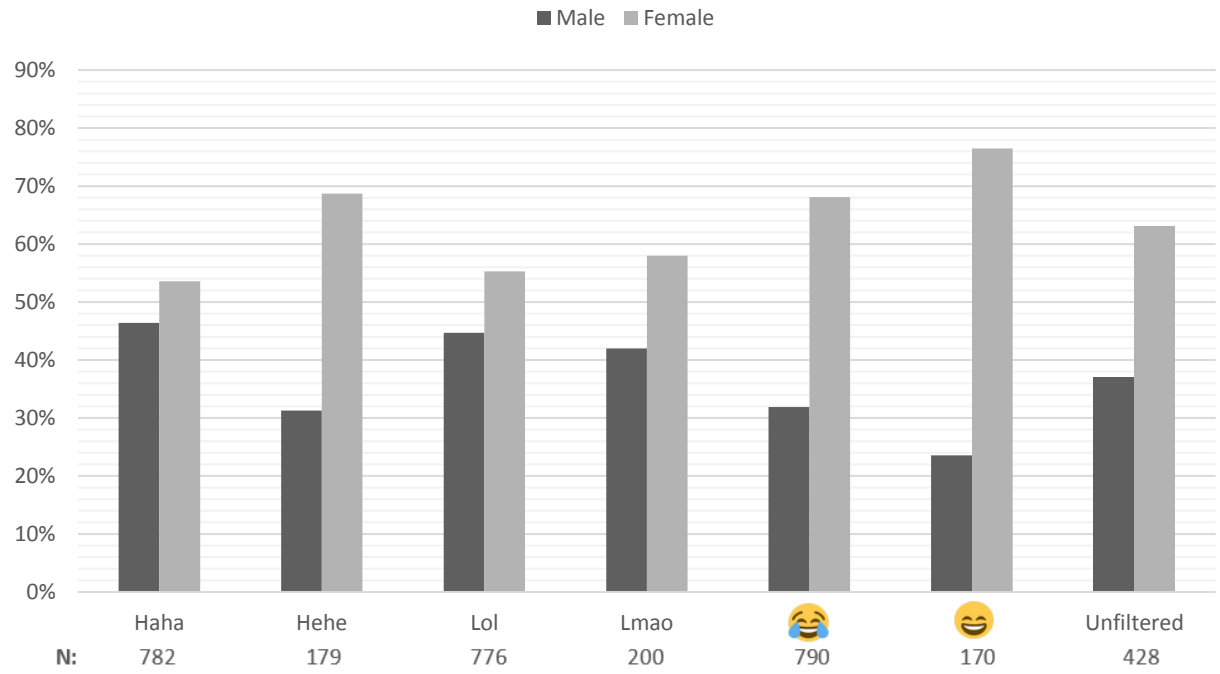
5.7.1: Gender of Tweeter

Most studies of gendered language on twitter have indicated that emoticons as well as "emotive abbreviations" and laughter words like *haha*, *hehe*, *lol*, and *lmao*, are much more commonly used by females than by males (Bamman et al., 2014). This goal of this analysis is to determine whether each of the forms of written laughter can be shown to occur in tweets composed by males or by females using the methodology described above. For each laugh form under consideration the percentage of gender-tagged tweets tweeted by males and by females is compared. These male/female tweeter distributions

for tweets containing each laugh form were then compared to the male/female tweeter distributions for tweets which were a part of the original unfiltered sample.

The results of this analysis, as presented in Figure 19, are somewhat surprising. Previous research would predict that each of the forms in question ought to be associated with female tweeters. Though the frequency with which tweets containing each laugh form were tweeted by females was higher than the frequency by males, in only a few cases (*hehe*, 😂, and 😄) was this frequency higher than the frequency with which females were found to tweet in general. This essentially means that though *lol*, *haha*, and *lmao*, were tweeted more often by females than by males, tweets containing these three forms were actually more likely than the average tweet to have been contributed by *males*. This means that three of the laugh forms in question actually seem to be somewhat associated with male, rather than female tweeters.

Figure 19: Tweeter Gender



There is, of course, another question which could be asked, namely, is written laughter in general a marker of female speech? The answer appears to be no. A chi-square test of independence was performed in order to examine the relation between the presence of written laughter in a tweet and the likelihood that the tweeter was female. The relation between these variables was not significant. Tweets containing written laughter are not significantly more likely to be contributed by females than the average tweet on the site $x^2(1, N=3325)=0.9821$, $p>0.1$.

It does seem worthy of note that two of the three forms which do appear to be associated with female tweeters are emoticon forms. A second chi-square test of independence was therefore preformed in order to examine the relation between the category of written laughter in a tweet (emoticon or spelled-out) and the likelihood that the tweeter was female. The relation between these variables *was* significant. Tweets containing emoticon laughter are significantly more likely to be contributed by females than are tweets containing spelled-out laughter forms $x^2(1, N=2897)=48.74$, $p<0.01$.

5.7.2: The Relationship of Gender to Offer/Acceptance Patterns

As has been discussed repeatedly above, face-to-face laughter generally takes on an offer/acceptance pattern in conversation. In addition to this, many studies have also found an association between the gender of conversational participants and their behavior with regard to laughter acceptance. These studies are discussed in the literature review of this paper, but in general, males are less likely to accept offered laughter, especially when that laughter is offered by a female, and are more likely to laugh when laughter is not offered, especially when the previous utterance was produced by a female. Contrastively females are more likely to accept offered laughter, especially when it is offered by a male, and less likely to laugh when laughter is not offered, especially when the previous speaker was male.

The examination of offer/acceptance patterns in Section 5.6.2 above found that the presence of spelled-out laugh forms in a tweet is associated with higher rates of laughter in previous and subsequent context tweets.[23] The position of laugh forms within the tweet, tweet-initial or tweet-final, was also found to be a relevant factor in the presence of spelled-out laughter in previous and subsequent tweets. Tweet-initial laughter was generally associated with spelled-out laughter in a previous tweet, while tweet-final laughter was generally associated with spelled-out laughter in a subsequent tweet.[24] Where the analyses preformed in Section 5.6.2.2 took the position of a laugh particle within a tweet to be the primary variable in whether or not context tweets contained laughter, this analysis considers the possibility that the gender of participants is that primary variable.

Because the analysis in 5.6.2.3 showed that the relationship between the presence of several of the laugh forms in question in dataset tweets and the presence of emoticon laughter in context tweets was small to non-existent, this analysis only examines the presence of spelled-out laughter in context tweets. Otherwise, the same tagging mechanisms were used to tag context tweets as containing spelled-out laughter as were used above.[25] Due to the amount of labor required in order to tag the gender of the tweeters of context tweets for this analysis, as well as the need for large datasets in order to obtain powerful statistics, this analysis only considered the forms *haha*, *lol*, and 😂.

The graphs in Figures 20 and 21 represent the results of this analysis. Figure 20 examines the influence that the gender of both the dataset tweeter and the previous tweeter has on the frequency of spelled-out laughter in previous tweets, while Figure 21 examines the influence that the gender of both the dataset tweeter and the subsequent tweeter has on the frequency of spelled-out laughter in

[23] Here we are once again discussing the frequency with the previous tweets of dataset tweets containing spelled-out laugh forms contained laughter. See the introduction of Section 5.6.2 for definitions of these terms.
[24] See Section 5.6.2.2
[25] See Section 5.6.2.2

subsequent tweets. N-values are low because frequencies are given in terms of, for example, the number of previous tweets tweeted by males before dataset tweets also tweeted by males.

Figure 20: Frequencies with which Previous Tweets Contained Spelled-out Laughter for Various Dataset Tweeter Gender and Context Tweet Gender Combinations

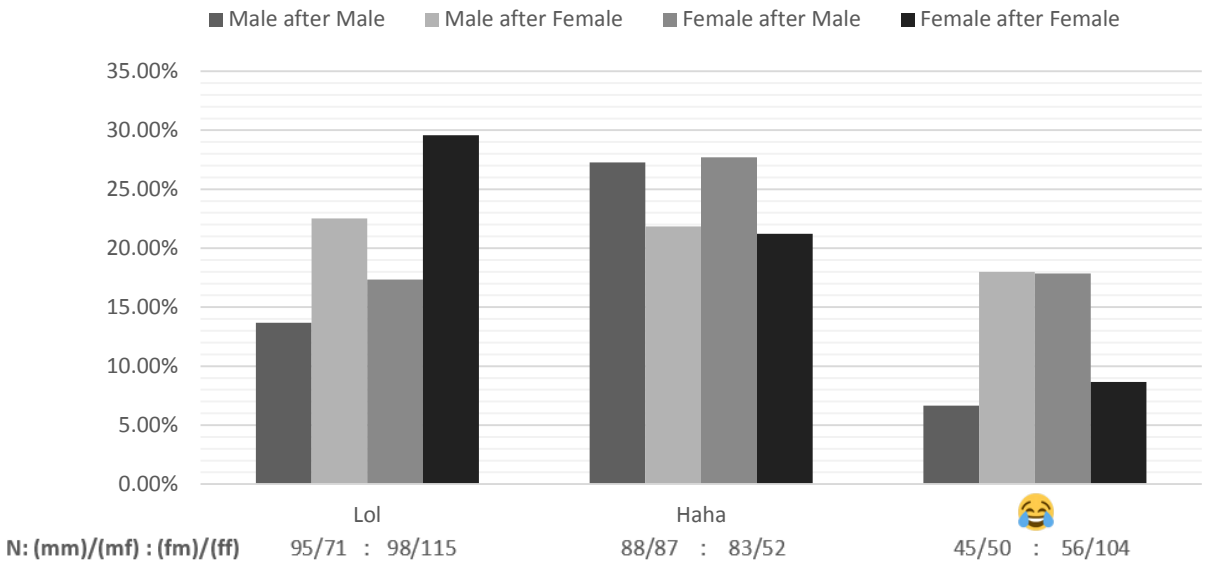


Figure 21: Frequencies with which Subsequent tweets Contained Spelled-out Laughter for Various Dataset Tweeter Gender and Context Tweeter Gender Combinations

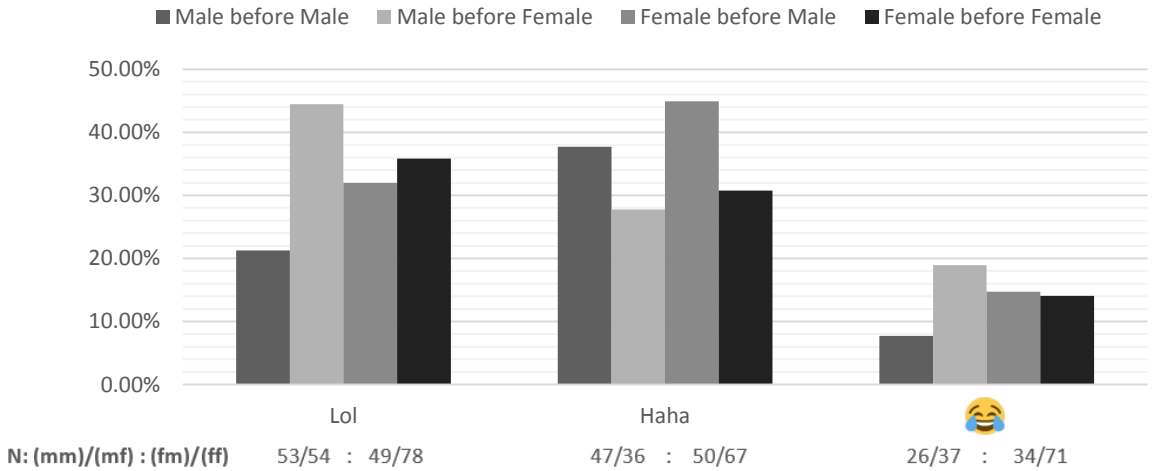


Figure 20 suggests that any influence that gender has on offer/acceptance patterns of written laughter is unrelated to what is seen in face-to-face laughter. Where in face-to-face communication males are unlikely to accept the laughter invitations of females, for at least two forms above (*lol* and 😂), male tweeters tend to use written laughter more often in response to the laughter of females than that of males. However, females do seem to show the expected pattern, at least for the forms *haha* and 😂. Females more frequently respond to the laughter of males than that of females. The exception to this trend is *lol*. Females seem more likely to respond to female laughter with *lol* than they are to respond to male laughter this way.

The patterns shown in Figure 21 resembled the patterning observed in Figure 20. Males using *lol* or 😂 were more likely to receive laughter acceptances from females. Females using *lol* or 😂 were about equally-likely to be answered by laughter from males as from females, while those using *haha* were more likely to receive a laughter response from males than from females.

Overall, this analysis revealed that there is no generalizable relationship between the genders of the dataset tweeter and the context tweeter and the likelihood that context tweets contained spelled-out laughter. It may be that gender has little influence on laughter acceptance rates because the gender of Twitter users is less physically obvious than in face-to-face conversation.

6: Discussion

This project had two overlapping goals. The primary was to establish preliminary information about the use of written laughter on Twitter, the second was to evaluate the extent to which we make use of our knowledge of physical laughter when writing laughter. In this section the analyses above will be examined as a whole with regard to the insights they can provide moving towards a better understanding of written laughter.

There appear to be three major types of written laughter, emoticon laughter, *lol*-type laughter (*lol*, *lmao*, and *hehe*), and *haha.* These are the same three groups which were first discussed in Figure 9 above. A table outlining the general behavior of each of these groups can be found in Table 2. [26]

Table 2: Similarities and differences between written laughter groups

#	Traits	Emoji Laughter	*Lol*-like Laughter	*Haha*
			Spelled-Out Forms	
1	The presence of this laugh form in a tweet increases the likelihood that a second laugh form will also occur in that tweet (5.2)	■	■	■
2	More likely to co-occur with spelled-out laugh forms (5.2)	■		
3	Tweets containing this laugh form are more likely to be specifically-targeted than are non-laughing tweets (5.3)	■	■	■
4	Tweets containing this laugh form are more likely to be specifically-targeted than not (5.3)		■	■
5	Form occurs initially in 20-30% of cases (5.4)		■	■
6	Tweets containing this laugh form tweet-initially are more likely to receive specific recipient design than those which contain this laugh form tweet-finally (5.5)	■	■	■
7	Tweet-final laughter is associated with specific recipient design (5.5)			■
8	Initial-position laughter is associated with the presence of externally-referent pronouns (5.6.1)		■	■
9	Tweets containing these laugh forms are more likely to contain spelled-out laughter in their context tweets than are tweets containing no laughter (5.6.2.1)		■	■
10	Tweets containing these laugh forms in tweet-initial position are more likely to contain laughter in previous tweets than tweets containing them finally (5.6.2.2)	■	■	
11	Tweets containing these laugh forms in tweet-final position are more likely to contain laughter in subsequent tweets than tweets containing initially (5.6.2.2)		■	■
12	Tweets containing these laugh forms are more likely to contain emoticon laughter in their context tweets than are tweets containing no laughter (5.6.2.3)	■		

[26] Shaded squares represent situations for which the association is observable

Though *lol*-type laughter does differ in some respects from *haha*-type laughter, the most striking differences can be drawn between emoji laughter and spelled-out laughter. In several cases, *lol*-type laughter and *haha*-type laughter behave similarly, while emoji laughter does not (see rows 2, 4, 5, 8, 9, 11, and 12). In general the spelled-out laugh forms behave in a way which is more reminiscent of spoken laughter than do the emoji laugh forms. They are more tightly associated with interpersonal exchanges and the locations in which they occur seem to be more strongly associated with previous and subsequent spelled laughter as well as with external reference. The lack of association between the presence of spelled-out laugh forms in a tweet and the presence of emoji in the previous and subsequent tweets (see Table 2) is strong evidence for the spelled-out/emoji distinction.

The lack of distinction between onomatopoeic laughter (*haha, hehe*) and abbreviated laughter (*lol, lmao*) is also worthy of note. Though *haha* has been placed in an independent category, *hehe*, another onomatopoeic form, behaved very similarly to *lol* or *lmao*. Several studies have examined netspeak abbreviations as if they were all one category. This study suggests that this is likely not the case. Netspeak abbreviations are often treated as one vague and somewhat mysterious class of lexical items. Here two of these netspeak abbreviations, *lol* and *lmao*, clearly behave a great deal like the other forms of spelled-out laughter under observations. It is their membership in the category "spelled-out laughter" which determines their use in conversation, not their membership in the category "netspeak abbreviations." Some studies have also attempted to treat forms like *lol* and *lmao* as emoticons. This also does not seem to be an appropriate treatment, as these spelled-out laugh forms do seem to be used in ways which clearly distinguish them from the emoticon forms.

These spelled-out laughter forms, at the very least, behave similarly enough to spoken laughter that it is likely that individuals who spell out their laughter are drawing on knowledge of spoken laughter. There are, however, some significant differences in usage patterns, particularly with regard to the offer-acceptance patterns of laughter. For the most part written laughter is associated with spelled-

out laughter in a previous or subsequent tweet. However, a number of the subtleties of face-to-face laughter seem not to translate to spelled-out laughter. For example, tweets containing written laughter contain laughter in previous and subsequent tweets at much lower rates than might be predicted were they regulated exactly as face-to-face laughter is. In addition, the genders of participants in twitter exchanges seems to have little to no influence on the rates at which written laughter is offered and accepted.

These differences might be explained in several ways. First, physical laughter is regulated almost unconsciously, whereas spelled-out laughter is placed more deliberately (Provine, 2001). The more unconscious portions of our laughter knowledge may, therefore, be put-aside when writing out laughter. Another explanation may be the relatively slow pacing of Twitter exchanges. Though often these conversations happen in real time, "real time" for typists is much slower than for speakers. Differences in behavior may also be due to the fact that the "laughing together" phenomenon briefly described in the literature review, cannot take place through text.

7: Future Research

This project is a very preliminary examination of an enormous internet phenomenon. Written laughter occurs all over the internet, not just on Twitter. It shows up in blogs, "memes"[27], text messaging, and emails, not to mention poetry, movie scripts, spoken language, or graffiti art. This project has examined the use of only six written laugh forms in only one internet communication

[27] I am referring here to the colloquial rather than the academic usage of the word "meme." In academic circles "meme" is used to describe ideas or trends which take off through a society (Blackmore, 1999). These are often the ideas which some describe as "viral." In the online community, however, this word refers to a specific sort of viral idea. These tend to be pictures, often screenshots from movies or well-known youtube videos, subtitled with humorous quips. Over time some pictures become associated with a certain sort of commentary and a certain sort of language use. *Lol*-cats are perhaps the most famous sort of "meme." Examples of "memes" can be found in Appendix 1.

environment. There are therefore a wealth of environments which could be the topics of future research.

First and foremost, this study was quantitative. Even utilizing this same data, a qualitative study could be conducted examining the ways in which written laughter is used in order to manipulate the flow of conversation or the meaning of utterances. A line of research like this might serve to fulfil the first half of the two pronged question posed in 4.1, it might serve to provide evidence that the intersubjective meaning of written laughter is similar to the intersubjective meaning of physical laugher. This is a topic which has been largely ignored in the present project which very much needs to be examined.

Another promising avenue for research would be to examine the use of written laughter in different SMS environments, such as text messaging, WhatsApp, Snapchat, and various instant messaging platforms. It would be interesting to see whether the results found here are repeatable in other similar environments or if laughter practices are specific to individual SMS platforms.

Similarly it would be interesting to see examine written laughter practices in longer-form messaging environments, such as email, Reddit, Tumblr, or the comment sections for media posts. To what degree might written laughter use in these environments mirror its use on Twitter? Moving forward, we might to look at the use of written laughter in non-messaging online environments like blogs, “meme” or lol-cat subtitling, and web comics. In this vein, many more traditional artistic genres such as poetry and script-writing have begun to make use of these online laughter words. Are written laugh forms used differently in more “formal” genres?

In addition to their online use, many written laugh forms have made a transition to spoken language as well. In these environments words like *lol* are able to co-exist with actual, physical laughter. It might be very enlightening to examine the similarities and difference in the usage of *lol* or *haha* in face-to-face interaction and in online interaction. If *lol* is being used in face-to-face interaction it must

mean something more than simple laughter. Looking at the ways in which each of these laugh words are used in face-to-face conversation could help to unravel the meaning of forms like *haha* or *lmao*, distinguishing between the meanings which approximate face-to-face laughter from additional meanings they take on as forms of written laughter.

This study also has revealed some interesting questions relevant not only to written laughter, but to netspeak forms like emoticons and abbreviations in general. There is a suggestion in these data that including laugh categories in general analyses for "netspeak abbreviations" may be misleading, as *lol* and *lmao* behaved so similarly to *haha* and especially *hehe*. It would be interesting to see whether the behavior of these two forms differs from the behavior of *omg*, *brb*, or *wtf*. This might be particularly interesting with regard to the recipient-design distributions of tweets containing these forms.

This study has similarly revealed some fairly strong differences in the behavior of spelled-out and emoticon laugh forms which deserve more extensive investigation. In particular, the relative lack of tweet-initial emoticons is interesting. Why are emoticons not used initially? Is what they do to the meaning and interpretation of a tweet different from what forms like *lol* or *haha* do in some more meaningful way? The intuitions given in section 3.1.6 of this paper about 😊 revealed that this form is only very weakly related to laughter in the minds of users, and yet its behavior through several comparisons was very similar to 😂's. Does that imply that these forms' identities as emoji are more important to their use and interpretation than their identities as laugh forms?

Additional phenomena of interest include rules governing the ways in which the orthography of various laugh forms can be varied in order to achieve different sorts of meaning. What is the difference between a capitalized laugh form and its lowercase counterpart (*lol* vs *LOL*)? What effect do repeated vowels have (*lol* vs *loooooool*)? What about the number of end-to-end repetitions of a laugh form (*lol* vs *lololol*)? Are these effects form-specific or universal? These and other aspects of orthographic variation such as surrounding punctuation and layout are worthy of future research.

The enormous online linguistic community which has developed on the internet over the last decade-and-a-half has far outstripped our academic understanding. There is a vast amount of data from numerous discourse environments available to linguists desiring to conduct their research online. Some linguists have a tendency to see this information as "just text," or, more often, as less essential to understanding the human linguistic system. But just as it is important to study Language from the perspective of more than just the English language, it is important to study Language from more than just the perspective of spoken language. This is why sign languages are studied so extensively, and it is why real-time written language should be studied as well.

Conclusion

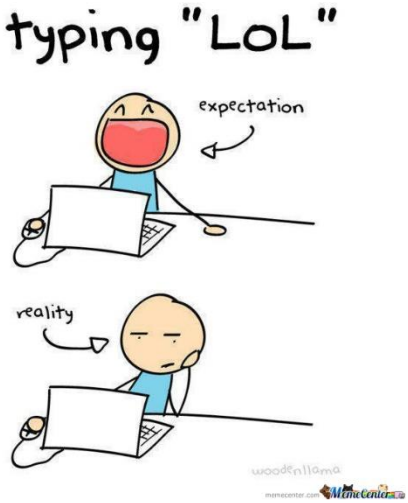
Whether it is sent over the telegraph wire or bounced off a satellite, spelled-out laughter is a fixture of SMS communications of all kinds. This project represents an initial attempt to understand why that might be. By examining data from the social networking site Twitter, some generalizations about different patterns of usage were drawn. Though any real understanding of written laughter is still in its nascent stages, this project did find that emoji forms of laughter behave fairly differently from spelled-out forms, but that abbreviated forms and onomatopoeic forms behave fairly similarly. Some of the patterns seen in the use of face-to-face laughter, such as its association with interpersonal exchanges and its invitation/acceptance patterns, were found to also hold for written laughter, while others, such as gendered patterning in invitation/acceptance and, in some cases the correlation between the location of a laugh within a turn and its identity as a laugh invitation or a laugh acceptance, could not be shown with significance. There is a long way to go in the study of written laughter, but it is my hope that this project will provide at least a small pool of information from which future projects may draw ideas.

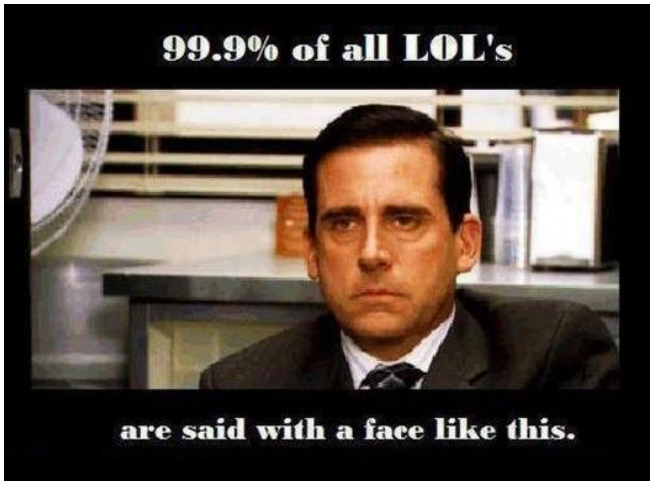
Acknowledgements:

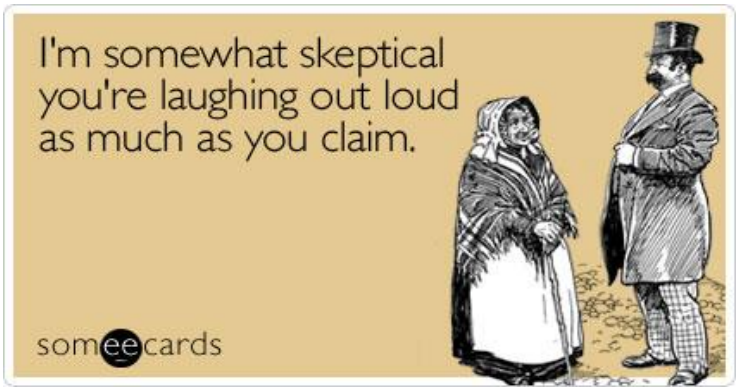
I would first and foremost like to thank my amazing thesis advisor, Debby Keller-Cohen, for all of the time and energy she's put into helping me through the various stages of this project. I couldn't have done it without her. I would also like to thank my second reader Robin Queen. I would also like to thank the other students who participated in the University of Michigan Honors Summer Fellows program over the summer of 2014, along with the program's organizers. Their help in brainstorming topics and in getting this project off the ground was invaluable.

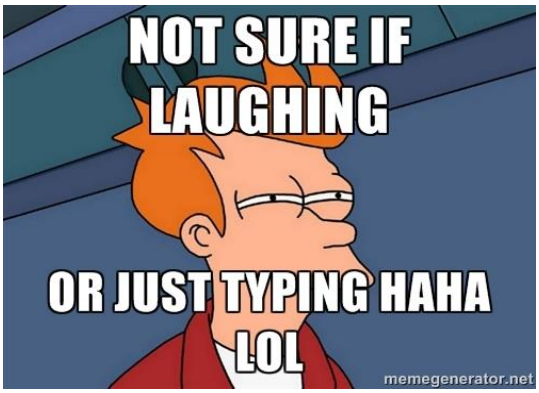
Appendices

Appendix 1: Observations about Written Laughter from Multimedia Humor









Appendix 2: Tweet-Initial Laughter References Previous/Ongoing Topic

Example 1:

-Lol-0882

Example 2:

-Lol-0883

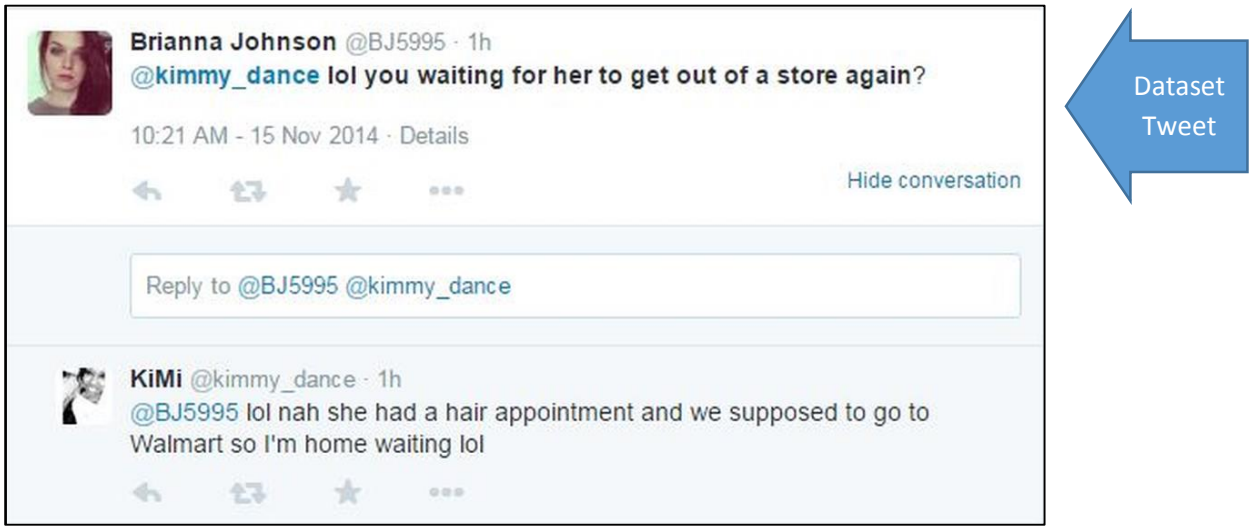
Example 3

-Haha-0604

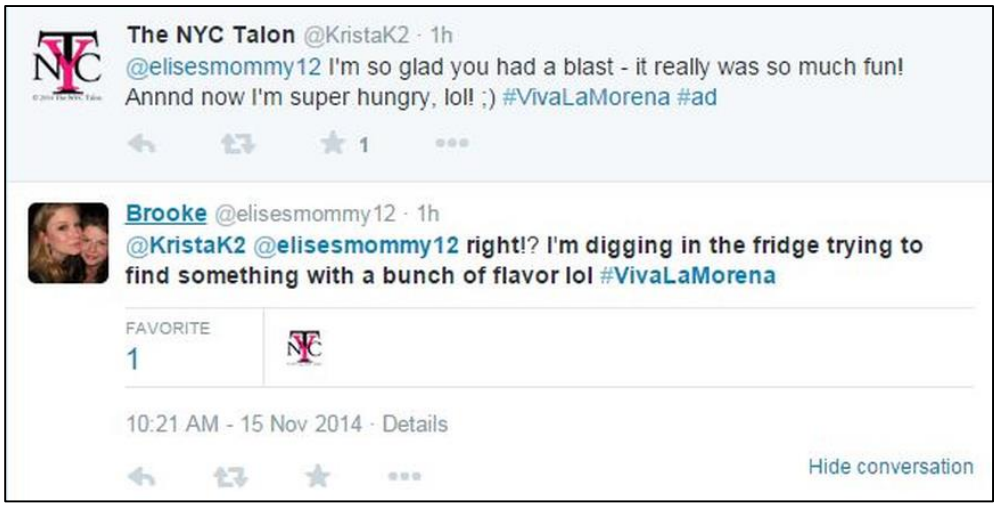
Appendix 3: Examples of Laughter in Previous and Subsequent Tweets for Various Laugh Locations:

Example 1: Tweet-Initial Laughter with a Previous Tweet containing Spelled-out Laughter (haha 241):

Example 2: Tweet-Initial Laughter with a Subsequent Tweet containing Spelled-out Laughter (lol 34):



Example 3: Tweet-Final Laughter with a Previous Tweet containing Emoticon Laughter (lol 15)



Example 4: Tweet-final Laughter with a Subsequent Tweet containing both Spelled-out and Emoticon Laughter (haha 36):

Works Cited

Alan. (2006). lmao. Retrieved from http://www.urbandictionary.com/define.php?term=lmao&page=9

Bamman, D., Eisenstein, J., & Schnoebelen, T. (2014). Gender Identity and Lexical Variation in Social Media. *Journal of Sociolinguistics, 18*(2). doi: 10.1111/josl.12080

Baron, D. (2011). Cyber-laugh: everything old is new again, haha. Retrieved from https://illinois.edu/blog/view/25/45228

Bateson, G. (1955). A Theory of Play and Fantasy *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution and Epistemology* (pp. 183-198). Northvale, New Jersey: Jason Aronson Inc.

Bennett, J. (2014, July 27). The Emoji Have Won the Battle of Words. *The New York Times*. Retrieved from http://www.nytimes.com/2014/07/27/fashion/emoji-have-won-the-battle-of-words.html?_r=0

Blackmore, S. (1999). *The Meme Machine*.

Brandon, J. (2008). FWIW -- The Origins of 'Net Shorthand. *PCWorld*. Retrieved from http://www.pcworld.com/article/153504/net_shorthand_origins.html

(2012, June 24). *That's What They Say* [Radio]. Retrieved from http://michiganradio.org/post/ever-wonder-what-lol-means

Derks, D., Bos, A. E. R., & Grumbkow, J. v. (2008). Emoticons and Online Message Interpretation. *Social Science Computer Review, 26*(3), 379-388.

Dresner, E., & Herring, S. C. (2010). Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force. *Communication Theory, 20*, 249-268. doi: 10.1111/j.1468-2885.2010.01362.x

dude, l. (2005). lol. *Urban Dictionary.* from http://www.urbandictionary.com/define.php?term=Lol&defid=1079038

Edel, L. (1989, May 8, 1989). MO_ICONS_PLEASE. *FidoNews: International FidoNet Association Newsletter*. Retrieved from http://www.textfiles.com/fidonet-on-the-internet/878889/fido0619.txt

Edmondson, M. S. (1987). Notes on Laughter. *Anthropological Linguistics, 29*, 23-34.

Entity1037. (2015). Haha. Retrieved from http://www.urbandictionary.com/define.php?term=haha&defid=8075298

EXPLIZIT. (2005). lol. *Urban Dictionary.* 2015, from http://www.urbandictionary.com/define.php?term=Lol

Face with Tears of Joy. (2015). *iEmoji.com.* from http://www.iemoji.com/view/emoji/26/people/face-with-tears-of-joy

Friends They Never Meet: Acquainteneces Made by the Telegraph Key. Confidences Exchanged between Men who have Never Seen Each Other - Their Peculiar Conversational Abbreviations. (1890, November 30, 1980). *New York Times*.

Glenn, P. J. (1989). Initiating Shared Laughter in Multi-Party Conversations. *Western Journal of Speech Communication, 53*(Spring), 127-149.

Glenn, P. J. (1991). Current Speaker Initiation of Two-Party Shared Laughter. *Research on Language and Social Interaction, 25*, 139-162.

Glenn, P. J. (2003). *Laughter in Interaction*. New York: Cambridge University Press.

Glenn, P. J. (2010). Interviewer Laughs: Shared Laughter and Asymmetries in Employment Interviews. *Journal of Pragmatics, 42*, 1485-1498.

Glenn, P. J., & Holt, E. (2013). Introduction. In P. J. Glenn & E. Holt (Eds.), *Studies of Laughter in Interaction* (pp. 1-22). New York: Bloomsbury.

Glenn, P. J., & Knapp, M. L. (1987). The Interactive Framing of Play in Adult Conversations. *Communication Quarterly, 35*(1), 48-66.
Goffman, E. (1974). *Frame Analysis: An Essay on the Organization of Experience*. Boston: Northeastern University Press.
guy12345. (2009). hehe. Retrieved from http://www.urbandictionary.com/define.php?term=hehe&defid=3904004
ha-ha. (2015).
Heaney, K. (2014a). The 42 Ways to Type Laughter, Define: Here's what your typed and texted LOL-ing REALLY means. Retrieved from http://www.buzzfeed.com/katieheaney/the-42-ways-to-type-laughter-defined#.shwKDjARK
Heaney, K. (2014b). The 58 Yellow Smiley Emoji, Defined. Retrieved from http://www.buzzfeed.com/katieheaney/the-58-yellow-smiley-emoji-defined#.wgBjRWoMj
Hiscott, R. (2014, 07/17/2014). LOL, WTF? The Origin Stories Of Your Favorite Internet Acronyms. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/2014/07/17/internet-acronyms_n_5585425.html
Hobbes, T. (1640). *The Elements of Law, Natural and Politic*. London: Unwin Brothers.
Holt, E. (2011). On the Nature of "Laughables": Laughter as a Response to Overdone Figurative Phrases. *Pragmatics, 21*(3), 393-410.
Holt, E. (2013). "There's Many a True Word Said in Jest": Seriousness and Nonseriousness in Interaction. In E. Holt & P. J. Glenn (Eds.), *Studies of Laughter in Interaction* (pp. 69-89). New York: Bloomsbury.
Jefferson, G. (1979). A Technique for Inviting Laughter and its Subsequent Acceptance Declination. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 79-96). New York: Irvington Publishers INC.
Jefferson, G. (1984). On the Organization of Laughter in Talk about Troubles. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 346-369). New York: Cambridge University Press.
Jefferson, G. (2004). A Note on Laughter in 'Male-Female' Interaction. *Discourse Studies, 6*(1), 117-133.
Jefferson, G., Sacks, H., & Schegloff, E. (1977). *Preliminary Notes on the Sequential Organization of Laughter*. Paper presented at the Informal Seminars of the Ethology-Hormones-Psychosomatics Research Groups, University of Pennsylvania.
kirshteeen. (2009). lmao. Retrieved from http://www.urbandictionary.com/define.php?term=lmao&defid=3688532
Krohn, F. B. (2004). A Generational Approach to Using Emoticons as Nonverbal Communication. *Journal of Technical Writing and Communication, 34*(4), 321-328.
Lampert, M. D., & Ervin-Tripp, S. M. (2006). Risky Laughter: Teasing and Self-directed Joking among Male and Female Friends. *Journal of Pragmatics, 38*, 51-72.
Lee, V., & Wagner, H. (2002). The Effect of Social Presence on the Facial and Verbal Expression of Emotion and the Interrelationships Among Emotion Components. *Journal of Nonverbal Behavior, 26*(1), 3-25.
Løkken, R. (2006). haha. Retrieved from http://www.urbandictionary.com/define.php?term=haha&defid=2043167
LOL, *int.* and *n.2*. (2015). Oxford University Press.
Lyddy, F., Farina, F., Hanney, J., Farrell, L., & O'Neil, N. K. (2014). An Analysis of Language in University Students' Text Messages. *Journal of Computer-Mediated Communication, 19*, 546-561. doi: 10.1111/jcc4.12045
McWhorter, J. (2013). Txtng is killing language. JK!!! *TED* [Video File].

Menzer, M. J. (2004). Ælfric's English "Grammar". *The Journal of English and Germanic Philology, 103*(1), 106-124.
monsieur_d. (2005). hehe. Retrieved from http://www.urbandictionary.com/define.php?term=hehe&defid=1466854
Morgan, J. (2011, April 8, 2011). Why did LOL infiltrate the language? *BBC News Magazine*.
Moss, C. (2014). Cosmopolitan Has A Guide On How To 'Decode' Men's Emoji. Retrieved from http://www.businessinsider.com/cosmopolitan-how-to-decode-his-emoji-2014-10
Neuendorf, K. A., & Fennell, T. (1988). A Social Facilitation View of the Generation of Humor and Mirth Reactions: Effects of a Laugh Track. *Central States Speech Journal, 39*(1), 37-48.
no_one_2000. (2005). lol. *Urban Dictionary.* Retrieved 3/9/2015, from http://www.urbandictionary.com/define.php?term=Lol&defid=1406597
Osborne, K., & Chapman, A. J. (1976). *Suppression of Adult Laughter: An Experimental Approach*. Paper presented at the International Conference on Humour and Laughter, Cardiff, Wales.
paperstars. (2009). hehe. Retrieved from http://www.urbandictionary.com/define.php?term=hehe&defid=3716398
Pearson, W. The Origin of LOL. 2015, from http://pages.cpsc.ucalgary.ca/~crwth/LOL.html
Peräkylä, A. (2007). Conversation Analysis *Blackwell Encyclopedia of Sociology Online*.
Provine, R. R. (2001). *Laughter: A Scientific Investigation*. United States: Penguin Books.
Provine, R. R., Spencer, R. J., & Mandell, D. L. (2007). Emotional Expression Online: Emoticons Punctuate Website Text Messages. *Journal of Language and Social Psychology, 26*(3), 299-307. doi: 10.1177/0261927X06303481
. Public API. (2013): Twitter.
Rees, C. E., & Monrouxe, L. V. (2010). "I should be lucky ha ha ha ha": The Construction of Power, Identity and Gender through Laughter within Medical Workplace Learning Encounters. *Journal of Pragmatics, 42*, 3384-3399.
rikochet. (2008). haha. Retrieved from http://www.urbandictionary.com/define.php?term=haha&defid=3248100
Rothenberg, M. (Producer). (2015). emojitracker. Retrieved from http://www.emojitracker.com/
Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(5), 696-735.
Sebastian. (2014). What your favorite emojis really mean (23 photos). Retrieved from http://thechive.com/2014/10/14/what-your-favorite-emojis-really-mean-23-photos/
Smith, R. E. (2011). Urban dictionary: youth slanguage and the redefining of definition. *English Today, 27*(4), 43-48.
The Streaming APIs: Documentation. (2014). Retrieved 11/9/2014, from https://dev.twitter.com/streaming/overview
TheDanishSugarfairy. (2008). lmao. *Urban Dictionary.* from http://www.urbandictionary.com/define.php?term=lmao&defid=3259576
Toole, C. (2014). Emoji Dictionary: The Honest Meanings Behind Your 15 Favorite Emojis. Retrieved from http://elitedaily.com/humor/emoji-dictonary-true-meanings-definitions/870508/
Vuolo, M. (2013). *LMAO*: A history. Retrieved from http://www.slate.com/blogs/lexicon_valley/2013/10/03/lmao_laughing_my_ass_off_has_a_long_lineage_that_dates_back_to_at_least.html
Walther, J. B. (2006). Nonverbal Dynamics in Computer-Mediated Communication or :(and the Net :('s with You, :) and You :) Alone. In V. Manusov & M. L. Patterson (Eds.), (pp. 461-480): Sage Publishers, Inc.
Walther, J. B., & D'Addario, K. P. (2001). The Impacts of Emoticons on Message Interpretation in Computer-Mediated Communication. *Social Science Computer Review, 19*.

Weller, K., Bruns, A., Burgess, J., Mahrt, M., & Puschmann, C. (Eds.). (2013). *Twitter and Society* (Vol. 89). New York.

Xy, L., Yi, C., & Xu, Y. (2007). *Emotional Expression Online: the Impact of Task, Relationship, and Personality Perception on Emoticon Usage in Instant Messenger*. Paper presented at the Pacific Asia Conference on Information Systems (PACIS), Auckland, New Zealand.

Yang, L., Sun, T., Zhang, M., & Mei, Q. (2012). *We know what @you #tag: does the dual role affect hashtag adoption?* Paper presented at the WWW '12: The 21st international conference on World Wide Web.